

AD-A171 140

STRUCTURING EXPERT JUDGMENT IN S/V ANALYSIS

Paula G. John
David A. Strobar
Gary A. Klein

Klein Associates, Inc
P.O. Box 264
Yellow Springs, Ohio 45387

April 1986

Final Report



DTIC
ELECTED
S JUL 14 1986
AUG. 22 D
D

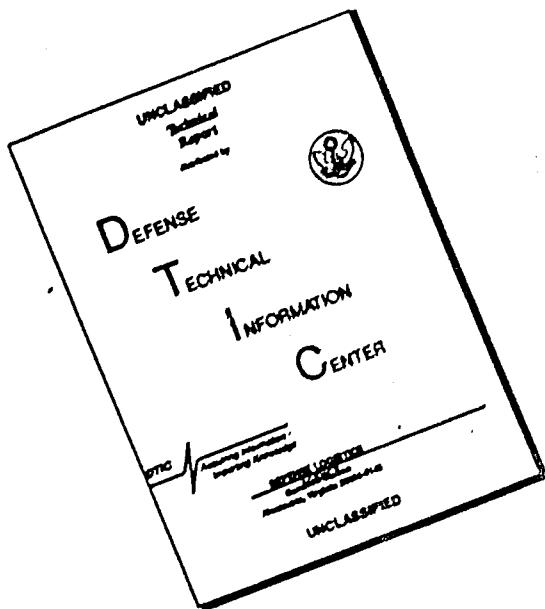
Approved for public release; distribution unlimited.

AMC FILE COPY

AIR FORCE WEAPONS LABORATORY
Air Force Systems Command
Kirtland Air Force Base, NM 87117-6008

86 8 21 2 088

DISCLAIMER NOTICE



**THIS DOCUMENT IS BEST
QUALITY AVAILABLE. THE COPY
FURNISHED TO DTIC CONTAINED
A SIGNIFICANT NUMBER OF
PAGES WHICH DO NOT
REPRODUCE LEGIBLY.**

This final report was prepared by Klein Associates, Inc, Yellow Springs, Ohio, under Contract F29601-85-C-0063, Job Order 3005NT11 with the Air Force Weapons Laboratory, Kirtland Air Force Base, New Mexico. Dr Timothy J. Ross (NTES) was the Laboratory Project Officer-in-Charge.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This report has been authored by a contractor of the United States Government. Accordingly, the United States Government retains a nonexclusive, royalty-free license to publish or reproduce the material contained herein, or allow others to do so, for the United States Government purposes.

This report has been reviewed by the Public Affairs Office and is releasable to the National Technical Information Services (NTIS). At NTIS, it will be available to the general public, including foreign nations.

If your address has changed, if you wish to be removed from our mailing list, or if your organization no longer employs the addressee, please notify AFWL/NTES, Kirtland AFB, NM 87117-6008 to help us maintain a current mailing list.

This technical report has been reviewed and is approved for publication.



DR. TIMOTHY J. ROSS
Project Officer



DAVID H. ARTMAN, JR.
Major, USAF
Chief, Applications Branch

FOR THE COMMANDER



CARL L. DAVIDSON
Colonel, USAF
Chief, Civil Engineering Research Div

DO NOT RETURN COPIES OF THIS REPORT UNLESS CONTRACTUAL OBLIGATIONS OR NOTICE ON A SPECIFIC DOCUMENT REQUIRES THAT IT BE RETURNED.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

ADA 171140

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS										
2a SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.										
2b DECLASSIFICATION/DOWNGRADING SCHEDULE												
4 PERFORMING ORGANIZATION REPORT NUMBER(S)		5. MONITORING ORGANIZATION REPORT NUMBER(S) AFWL-TN-86-07										
6a NAME OF PERFORMING ORGANIZATION Klein Associates, Inc	6b. OFFICE SYMBOL <i>(If applicable)</i>	7a. NAME OF MONITORING ORGANIZATION Air Force Weapons Laboratory										
8c ADDRESS (City, State and ZIP Code) P.O. Box 264 Yellow Springs, OH 45387		7b. ADDRESS (City, State and ZIP Code) Kirtland Air Force Base, NM 87117-6008										
9a NAME OF FUNDING/SPONSORING ORGANIZATION	9b. OFFICE SYMBOL <i>(If applicable)</i>	10. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER F29601-85-C-0063										
11c ADDRESS (City, State and ZIP Code)		10. SOURCE OF FUNDING NOS. <table border="1"><tr><th>PROGRAM ELEMENT NO.</th><th>PROJECT NO.</th><th>TASK NO.</th><th>WORK UNIT NO.</th></tr><tr><td>65502F</td><td>3005</td><td>NT</td><td>11</td></tr></table>		PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.	65502F	3005	NT	11	
PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.									
65502F	3005	NT	11									
11 TITLE (Include Security Classification) STRUCTURING EXPERT JUDGMENT IN S/V ANALYSIS												
12 PERSONAL AUTHORISI John, Paula G.; Strobhar, David A.; and Klein, Gary A.												
13a TYPE OF REPORT Final	13b. TIME COVERED FROM Jun 85 to Jan 86	14 DATE OF REPORT (Yr., Mo., Day) 1986 April	15. PAGE COUNT 64									
16. SUPPLEMENTARY NOTATION												
17. COSATI CODES <table border="1"><tr><th>FIELD</th><th>GROUP</th><th>SUB. GR.</th></tr><tr><td>13</td><td>13</td><td></td></tr><tr><td>05</td><td>10</td><td></td></tr></table>	FIELD	GROUP	SUB. GR.	13	13		05	10		18. SUBJECT TERMS (Consume on reverse if necessary and identify by block number) Survivability Buried structures Comparison-based method Vulnerability Blast Knowledge engineering Expert opinions Judgment		
FIELD	GROUP	SUB. GR.										
13	13											
05	10											
19. ABSTRACT (Consume on reverse if necessary and identify by block number) <p>It is essential that the Air Force be able to predict the Survivability and Vulnerability (S/V) of structures exposed to various types of blast and shock waves. Direct testing is time-consuming, expensive, and often not feasible. Attempts to develop formal models have been limited because of the complexity of the problem. Therefore, the predictions are typically based on expert judgments. However, expert judgments can be biased and unreliable unless carefully structured.</p>												
This project examined the feasibility of using a Comparison-Based Prediction (CBP) method for structuring S/V judgments. The CBP method is designed to improve the quality of predictions by anchoring them in existing data and by creating an audit trail documenting how the predictions were made. A total of 31 Subject-Matter Experts (SMEs) were interviewed in an experimental paradigm contrasting predictions made using the CBP method with predictions made without using CBP. The research design was modified several times to (over)												
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED <input checked="" type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS <input type="checkbox"/>		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED										
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr Timothy J. Ross		22b. TELEPHONE NUMBER <i>(Include area code)</i> (505) 844-9087	22c. OFFICE SYMBOL NTES									

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE

19. ABSTRACT (Concluded)

incorporate improvements in the way CBP was applied to this domain, as well as to improve the experimental paradigm.

Results showed that the CBP method was feasible for use with these SMEs, and that it supported their normal but unstructured use of analogical reasoning. Further, when SMEs used the CBP method, the reliability of their predictions increased significantly: The standard deviations of the predicted vulnerability measures were reduced by over 30%. Finally, the CBP method documented the bases for the SMEs predictions, thereby serving as a knowledge elicitation tool for describing their judgments.

These results have implications for knowledge elicitation of S/V judgments in general, and for development of expert systems. Additionally, the results suggest the value of CBP for improving S/V predictions directly. Additional research is needed to study the predictive validity of the CBP method.

ACKNOWLEDGEMENTS

This study was done with funds awarded under the Defense Small Business Innovation Research Program, mandated under the Public Law 97-219.

The authors wish to thank Dr. Timothy Ross, our Contract Monitor at the Air Force Weapons Laboratory, for his help and support throughout this project. We are grateful also to the following firms and agencies who permitted personnel to be interviewed for this study:

ACTA
Aerospace Corporation
Air Force Weapons Laboratory
Applied Research Associates, Inc.
Army Corps of Engineers/Waterways Experiment Station
BDM Corporation
Engineering Mechanics Associates
Jaycor
Karagozian and Case
SAIC
Sandia National Laboratory
TRW
University of Illinois
University of New Mexico
Weidlinger Associates

Accession For	
NTIS	CRA&I
DTIC	TAB
Unannounced	
Justification	
By _____	
Distribution / _____	
Availability Codes	
Dist	Avail and/or Special
A-1	



TABLE OF CONTENTS

ABSTRACT	1
INTRODUCTION	2
S/V ASSESSMENT	2
EXPERT JUDGMENT	3
COMPARISON-BASED PREDICTION	4
OBJECTIVES	8
REVISED OBJECTIVES.	8
METHODS	9
PILOT STUDY	9
PHASE I STUDY DESIGN	9
PHASE II STUDY DESIGN	9
SUBJECTS	10
COMPARISON-BASED PREDICTION DESIGN	10
FINAL CBP STRATEGY	12
RESULTS	12
PHASE I NUMERICAL RESULTS	15
PHASE II NUMERICAL RESULTS	15
INFLUENCE OF SYSTEM PARAMETERS	16
SUBJECTIVE DATA	16
DISCUSSION	24
CONCLUSIONS AND RECOMMENDATIONS	26
REFERENCES	29
APPENDIX A: INTERVIEW MATERIALS	31
APPENDIX B: PREDICTION DATA	49
APPENDIX C: INITIAL CBP STRATEGY	53

FIGURES AND TABLES

Table 1: CBP Method	6
Table 2: Background of Subjects	11
Table 3: Phase I Results: Prediction Scores	13
Table 4: Phase II Results: Prediction Scores	14
Table 5: Causal Factor Emphasis: Case A2	17
Table 6: Causal Factor Emphasis: Case A3	18
Figure 1: Separate and Combined Failure Mode Estimates	20
Figure 2: Contrast of Linear vs. Non-Linear Assumptions on Estimated Degree of Damage	23
Figure A:1: SME Background Data Sheet	32
Figure A:2: Interview Summary Sheet	33
Figure A:3: Descriptions of Target Cases (As)	37
Figure A:4: Description of Comparison Cases (Bs). . .	42
Table B:1: Prediction Scores of Phase I SMEs	50
Table B:2: Prediction Scores of Phase II SMEs	51
Table C:1: Prediction Scores of Pilot Study SMEs . .	55

1.0 ABSTRACT

It is essential that the Air Force be able to predict the Survivability and Vulnerability (S/V) of structures exposed to various types of blast and shock waves. Direct testing is time-consuming, expensive, and often not feasible. Attempts to develop formal models have been limited because of the complexity of the problem. Therefore the predictions are typically based on expert judgments. However, expert judgments can be biased and unreliable unless carefully structured.

This project examined the feasibility of using a Comparison-Based Prediction (CBP) method for structuring S/V judgments. The CBP method is designed to improve the quality of predictions by anchoring them in existing data and by creating an audit trail documenting how the predictions were made. A total of thirty-one Subject-Matter Experts (SMEs) were interviewed in an experimental paradigm contrasting predictions made using the CBP method with predictions made without using CBP. The research design was modified several times to incorporate improvements in the way CBP was applied to this domain, as well as to improve the experimental paradigm.

Results showed that the CBP method was feasible for use with these SMEs, and that it supported their normal but unstructured use of analogical reasoning. Further, when SMEs used the CBP method, the reliability of their predictions increased significantly: the standard deviations of the predicted vulnerability measures were reduced by over 30%. Finally, the CBP method documented the bases for the SMEs predictions, thereby serving as a knowledge elicitation tool for describing their judgments.

These results have implications for knowledge elicitation of S/V judgments in general, and for development of expert systems. Additionally, the results suggest the value of CBP for improving S/V predictions directly. Additional research is needed to study the predictive validity of the CBP method.

2.0 INTRODUCTION

2.1 SURVIVABILITY/VULNERABILITY ASSESSMENT

Survivability/Vulnerability (S/V) assessment is the determination of a system's ability to continue to function following an external stressor. In a military environment, the external stressor is an attack with conventional, nuclear, or directed-energy weapons. The military systems include both above and below ground structures, vehicles, and equipment.

The problem of assessing the Survivability/Vulnerability of systems is a difficult process for a variety of reasons: lack of agreement on definitions, cost of testing, number and magnitude of system variables, and lack of an extensive empirical data base. The S/V assessment problem is further compounded when considering nuclear weapon effects as the system, by law, can be only simulated or approximated.

"Continue to function" (survivability) or "fail to function" (vulnerability) can mean different things to different people, depending upon what they think is the function of the system. Identical system responses can be interpreted in opposite ways. Breach of structural integrity and/or content failure immediately following the stressor can be used to define vulnerability. However, it can be argued that maintenance of the structure/equipment or time-to-restore-function would be a better metric of S/V. Minor roof damage can mean a machine shop survives but a computer center fails.

Survivability and vulnerability can also be used to describe different approaches to the assessment problem. New structures are designed to insure the survivability of the system. Existing systems are evaluated relative to their vulnerability. The terms survivability and vulnerability may therefore also refer to the goals (design vs. assessment) of the problem.

Testing of system responses is not only technically difficult but also extremely costly. As nuclear weapons cannot be used, conventional explosives must be designed that simulate nuclear weapon effects. Simplified scale models must be used due to the size of actual systems and simulation associated constraints. Instrumentation for the testing is difficult due to the extreme environmental conditions present. The cost of the tests is increased by the high quality control requirements necessary for scale model production.

The large size of the S/V problem hampers attempts to generate meaningful results. The S/V problem is not really one problem, but four problems: (1) blast simulation, (2) geological response, (3) structural response, and (4) content response. Each technical area contains numerous interacting variables which must be taken into consideration.

Blast simulation was not begun in earnest until after 1964. The Nuclear Test Ban Treaty necessitated that conventional explosives be designed that simulate nuclear blasts. This required the development of explosions with high peak pressures and rapid decay time.

Geological response has the largest data base. The cratering following nuclear blasts and the attenuation of pressure waves by the soil have been studied more than any other aspects of the S/V problem. But, because of the heterogeneous and anisotropic nature of soils, geological response still contains large uncertainties.

Structure and content response is data limited. Testing of structures, particularly to failure, is a recent practice. The utility of content response data must lag behind the collection of structural data, as the latter is a prerequisite to content analysis. Also, contents likely to require S/V analysis are changing in nature very rapidly. Thus rapid advances in electronics technology have resulted in greater gaps in the availability of relevant data about contents response.

There does not exist a comprehensive data base for S/V problems. Data quantity and quality appear to decrease as the problem progresses from initiation (i.e., blast) to conclusion (i.e., content response). Little or no real world data exist to augment test data. The paucity of data in key areas and the lack of a mechanism to integrate existing data hamper efforts to develop valid analytic system models.

In place of "hard" data, expert judgment or prediction must fill the gap. Sealed test data and analytic analysis must be synthesized into an assessment of system S/V. The final assessment is therefore not directly obtained from tests or analyses, but altered by expert judgment to account for assumptions and simplifications required by the tests and analyses.

A recent study of the USAF Weapons Laboratory indicated some of the problems in expert judgment relative to S/V. A group of experts were given photographs of tested structures and asked to describe (1) the degree of damage independent of function and (2) the mode of failure. Despite the instruction to ignore function, most experts based the degree of damage on some stated function. In addition, the experts in many cases disagreed on both degree of damage and failure mode. The study elicited excellent information on the assumptions of experts and the variables they consider. However, it pointed out the low reliability of unstructured expert judgment (Ross & Wong, 1985).

2.2 EXPERT JUDGMENT

The use of expert judgment presents a range of problems. Who is an expert? With a problem as complex as S/V analysis, experts must be drawn from many fields to incorporate different perspectives and information needs.

How reliable is the judgment? Experts do not always document the basis for their judgment, nor are they always able to do so. Expertise, after all, often seems to lie in subtle processes that the possessor cannot explain. For example, perceptual discriminations learned through great amounts of practice are not easy to articulate. It takes the work of trained knowledge engineers to elicit the description of such skills. They are not very accessible to the experts themselves, who may know the content area but not the psychological underpinnings of their own knowledge.

How can judgments be quantified for inclusion in a formal model, or even for comparison with other judgments? Experts often are given problems and provide solutions in their own framework, resulting in responses that seem to conflict with judgments from other experts working within other frameworks.

Klein Associates has conducted both basic and applied research in the psychology of expert judgment and decision making, and has developed and refined models of these processes. From these models we have developed a methodology for structuring expert judgments so that they can be more reliable as prediction tools, can be compared from one expert to another, and can be analyzed and evaluated. The method has proved feasible in areas where there are ambiguous or missing data or uncertainty about critical elements of the prediction scenario.

The current study was undertaken to test the feasibility of the method, called Comparison-Based Prediction (CBP), for application to the S/V assessment problem. Specifically, it was to be tested as a means for structuring the subjective element in the prediction process. The objective was to develop a method for gathering S/V predictions based on structured expert judgments. This method would have use in generating predictions, in providing inputs to other analytical methods or systems, and in developing a model of the S/V judgment process.

2.3 COMPARISON-BASED PREDICTION

Comparison-Based Prediction (CBP) is a methodology for making predictions when there are unknown parameters, missing data, or unclear objectives. Operationally, it is a way of structuring the judgments that experts make when they are called on to estimate unknown properties of a new situation. Formally, it is a system of reasoning by analogy, predicting to an unknown case by using what is known about a comparable case.

Reasoning by analogy is a natural process that contrasts with Bayesian and statistical decision models, which are hard to apply to most operational contexts.

A common example of this use of analogy comes from real estate. A realtor sets a price for a property, not by using a formal model and calculating all the variables, but by choosing a comparable sale and adjusting its price on the basis of differences between the two properties. Engineers have traditionally made use of analogies in prediction and design. They typically look for structural comparison. If their task is to predict how reliable a new piece of equipment is going to be, engineers use historical data for a basis of estimate.

A formalized version, called Comparability Analysis, is found in the Air Force (Tetmeyer, 1976). Developed in 1971, it is a way of explicitly using historical data to predict equipment reliability for the purposes of spare parts purchasing, manpower need projections, downtime forecasts, etc. Working directly from Air Force maintenance data, the engineer identifies a craft comparable to the one being planned. The next step is to derive an adjustment factor that reflects the differences between the comparison case and the new equipment. The third step is to present the rationale for the

adjustment factor. Next is to collect the operational data for the comparison system, showing how reliable that equipment has been under operational conditions. The last step is to adjust these historical data to generate a prediction.

In the decade since Comparability Analysis was developed, it has been applied in the Air Force to a variety of new aircraft. Each involved a variety of subsystems, so that hundreds of these studies have been conducted. Widenhouse and Romans (1977) collected evaluation data contrasting predictions with observed data. We analyzed these data (Klein & Gordon, 1984) by calculating Pearson product-moment correlations between predicted and actual time measures. For mean-time between failures, the correlation between predicted and actual data was a strong .76; for maintenance man-hours per flying-hours, the correlation was a high .84.

Klein Associates assessed the process of Comparability Analysis (Klein & Weitzenfeld, 1982) and presented an explanation of the logic underlying the use of comparison cases to derive predictions (Weitzenfeld, 1984; Weitzenfeld & Klein, 1982). We were interested in proving the method and in increasing its range of application beyond reliability and logistics. We studied three existing models of analogical reasoning, and found that none seemed to reflect the important aspects of Comparability Analysis: choosing an appropriate analogous situation; assessing the difference between it and the situation under study; and deriving an inference (prediction) by adjusting data obtained from the analogue.

We have suggested a model of analogical reasoning that emphasizes the role of causal factors (Klein, 1982; Weitzenfeld, 1984). This model states that for Situation A there is a set of causal factors ($x, y, z\dots$) that will determine or influence $T(A)$, the target characteristic of A to be estimated. Situation A could be a new aircraft duct system; causal factors x, y, and z could be the size of the aircraft, the material used, and a particular construction technique; and $T(A)$ could be the reliability of the system as measured by Mean-Time Between Failures.

In determining the target value, $T(A)$, we usually cannot identify all of the causal factors involved, their effects and interactions. Instead, an analogous situation or comparison case (Situation B, another duct system) is identified which reflects the same determinants as the target case. That is, for aircraft B, the same causal factors ($x, y, z\dots$) determine a corresponding value, $T(B)$, as a measure of system B.

Although the same causal factors affect both $T(A)$ and $T(B)$, it is unlikely that the values of the causal factors will be the same in both cases. In using $T(B)$ as an estimate of $T(A)$ we can note the differences in the values of each of the causal factors and make adjustments in our predictions to take these differences into account. Although checklists of causal factors can be provided, the method requires experts to use their experience in identifying the most important causal factors to use.

Comparison-Based Prediction is the methodology following from this model. The general CBP strategy (outlined in Table 1) begins with the definition of the target measure, T, and the identification of major determining (causal) factors known to affect it. Next a selection of possible comparison cases is identified.

From these, Subject Matter Experts (SMEs) choose one case, based on the similarity of the effect of the causal factors between it and the target case. The comparison case value that is analogous to the target case value $T(A)$ is specified as $T(B)$. SMEs then make a rough estimate of the differences expected between $T(B)$ and $T(A)$, most often only a judgment of whether $T(A)$ will be greater or less than $T(B)$.

Table 1
THE CBP METHOD

Setting up the Problem	- Stating the Problem in definitive terms: $T(A)$ = the Prediction Target
	Framing the questions: the Causal Factors and Prediction Scenario
Selecting the Resources	- Choosing the Comparison Case and $T(B)$ data - Choosing the SME(s)--Subject Matter Experts
Collecting and Analyzing the Data	- Interviewing the SME(s) Analyzing Causal Factors to obtain $T(A)$ from $T(B)$
Documenting the Process	- Recording the Process - Leaving an Audit Trail for others to follow

The SMEs then are guided through an examination of the effect of the expected differences in values of causal factors, until this effect can be quantified so as to produce an "applicable adjustment factor." This factor is then applied to operational data for $T(B)$, to yield a prediction for $T(A)$. Analysis of the differences among factors produced by SMEs can produce a confidence range for the prediction. The process is documented to provide an audit trail, so that the basis for the prediction can be understood and the findings adjusted should changes be made in the target case.

The CBP technique relies on the use of Subject Matter Experts (SMEs) who are knowledgeable about the domain of interest, in order to select optimal comparison cases and identify the relevant causal factors. The CBP approach elicits SME judgments through the use of a carefully structured interview with a format reflecting the general CBP process outlined in Table 1. The approach is data driven since the SMEs are generating

adjustment of operational data and giving their reasons for making these adjustments. There may be cases where no operational data are available. It is possible to proceed with a CBP approach by having the SMEs estimate the operational data, but this is not the ideal application of CBP method, and will reduce confidence in the outputs. However, this is often the state of affairs for the predictions where CBP is used, since this is usually the clearest situation where there are no alternative prediction methods.

An important element in the CBP strategy that can increase our confidence in the prediction is the development of an audit trail. The audit trail consists of a detailed description of the causal factors considered by the SME, and the impact estimated for each. By having an explicit set of causal factors to consider in determining adjustments, the SME has a set of concepts to use in posing the differences between the target case and the comparison case(s). This facilitates communication among SMEs and helps to standardize the variables considered in the prediction process. In addition, if the prediction is found to be inaccurate once operational data are obtained for the target case, the audit trail provides an opportunity to go back and see which considerations (causal factors) were responsible for the misjudgment. This process is obviously not possible when only an unstructured expert opinion has been obtained, that is, one which follows no defining guidelines and cannot be replicated.

It can be seen that CBP has several advantages over traditional prediction techniques. The CBP strategy is relatively easy and straightforward, and can be used even when there are unknown parameters, missing data, or unclear objectives. In addition, it requires relative judgments from the SME (evaluating one situation in relation to another), which seem easier for them to make than absolute judgments. Perhaps the most important strength of the method is that it grounds the predictions in established experience. Additionally, CBP creates an audit trail of the prediction process, which can later be used to evaluate and improve the prediction. Finally, it has high face validity in that it seems to be a structured form of a naturally occurring inference process, reasoning by analogy.

One brief validation study has been conducted to date. CBP was used to predict the outcome of an experiment on differences in effectiveness of functional and physical fidelity of training devices. Correlation of CBP predictions with test results was .90, accounting for 81% of the variance (Klein, 1986).

A major limitation of Comparison-Based Prediction is that it requires data about specific cases, not merely statistics about groups of cases or data such as would be required to support finite element analysis. This data base is not always available but can usually be estimated satisfactorily by SMEs. This limitation can, however, also be viewed as an advantage: it combines "real world" data with expert judgment, and this combination may produce more comprehensive estimates by forcing consideration of the varieties of response possible.

A strength of Comparison-Based Prediction is that it supports the natural analogical reasoning process that experts report they use when

required to give a subjective judgment. This value is twofold. First, it makes the judgments of several experts capable of direct evaluation and comparison because they were made within the same structural framework. Second, it provides a structure for capturing the judgment process so that it can be replicated and modeled for applications such as the development of expert systems. In problems such as S/V analysis, where subjective judgments are but one element of a formal model that contains much hard data, this factor may be of more importance than the power of CBP as a prediction technique.

3.0 OBJECTIVES

The original goal of the project had been stated in the proposal as "to develop a method for the prediction of Survivability/Vulnerability (S/V) to nuclear blast and shock to protective structural facilities. The particular problem is the characterization of expert judgment and technical intuition as elements in the assessment process."

There were two technical objectives for the study. The primary objective was to determine the feasibility and value of using a CBP strategy to estimate the nuclear survivability and vulnerability of protective structural facilities.

The secondary objective was to apply CBP to a current AFWL prediction problem.

In support of this secondary objective, the original design for this study was a simple controlled experiment, whereby predictions made by experts using CBP would be compared with those made by experts in the traditional unstructured way, and both sets of predictions would be validated against test data, from either empirical tests or simulations. Because no appropriate prediction problems were available -- that is, AFWL had no current prediction problems for study nor were any test data available as validation for a prediction study -- a new design had to be constructed, along with revised technical objectives.

3.1 REVISED OBJECTIVES

Since it would not be possible to validate CBP as a method for generating S/V predictions, we reframed the study, in consultation with the Contract Monitor, to test the feasibility of using CBP as a knowledge elicitation tool in S/V analysis. AFWL was exploring approaches to building an expert system model in this domain, ultimately to construct an artificial intelligence application. We therefore planned to test the feasibility of using CBP to extract from Subject Matter Experts (SMEs) the cognitive processes by which they made their predictions as well as the specific factors they considered. If CBP were proved successful in this effort, the knowledge so gained could become the basis for the expert system model.

4.0 METHODS

The description of methodology for this project must cover two topics, not only the experimental design for the study but also the application of the Comparison-Based Prediction technique. CBP itself as a method for prediction requires testing and adaptation to the problem under study. In this case, the CBP format was revised several times before a satisfactory application was achieved. The experimental design was made final only after the CBP technique was refined.

4.1 PILOT STUDY

The design was made final only after a small pilot study. This initial set of interviews, with four subjects at the University of Illinois, was used to refine the CBP presentation, clarify the prediction target definition, test the CBP strategy choice, verify the selection of causal factors, and generally examine the interview design.

4.2 PHASE I STUDY DESIGN

The study design began with the construction of five sample S/V prediction problems. At the start of the interview, each subject was presented these five prediction problems and asked to make a simple expert judgment of the vulnerability of each. These predictions were set aside while the study was introduced and the CBP methodology explained. The subject was then led through the CBP process for one of those five problems, the target case, producing a second prediction for it. He was then shown his initial prediction for that problem, and asked to evaluate which of the two he thought was more reliable.

Subjects were then allowed to revise their initial predictions for the other four problem cases. Differences between initial and final predictions for these non-target cases would indicate the influence of the CBP structure. Thus the study was to assess not the validity of CBP-structured judgments vs. other judgment modes, but rather the utility of CBP for supporting or enhancing the judgment process.

Each interview thus yielded ten scores: five predictions pre-CBP, one CBP-structured prediction, and four predictions post-CBP. These post-CBP predictions were revisions of initial predictions, not structured applications of the CBP technique.

4.3 PHASE II STUDY DESIGN

A preliminary analysis of data was made after fifteen Phase I interviews using this format. It was recognized that the initial prediction exercise, in which the subject made unstructured expert judgments for five cases, had itself become a CBP trial: subjects were using analogical reasoning already at this stage, making their initial judgments by comparing one case with another and adjusting their predictions for each on the basis of relative assessments. Thus the power of CBP as a structuring methodology would go unmeasured, since some of its effects would be already incorporated into the preliminary predictions.

The format was revised to correct for this condition and another twelve interviews were held. In these trials, subjects were given only two cases for which to make initial predictions, and then were led through the CBP process to make a structured judgment on one of them. They were then permitted to revise their initial judgment for the other, non-target case. (In order to reduce the effect of analogical reasoning used in the initial judgments, the case that would become the target for CBP prediction was presented second in the preliminary prediction stage. Thus any effect of preliminary reasoning by analogy would be already reflected in the pre-CBP score.) These interviews yielded four scores each: two predictions pre-CBP, one prediction using CBP structure, and one post-CBP prediction.

4.4 SUBJECTS

Subjects were chosen from a list suggested by AFWL of engineers with specialties in relevant areas both in government agencies and at contractor firms and universities. In some instances, experts thus chosen recommended colleagues for additional interviews. All had expertise in some area of vulnerability prediction. Thirty-one interviews were held, in six cities, yielding twenty-seven CBP-structured predictions. Four interviews were for the pilot study, fifteen for Phase I, and twelve for Phase II. Data describing the subjects are listed in Table 2.

4.5 COMPARISON-BASED PREDICTION DESIGN

Since this study was the first application of CBP to problems of S/V analysis, each element of the method was defined at a simplified level.

4.5.1 Target case. Two buried structures were hypothesized, one with an arched roof and one with a flat roof. The contents were specified as a commercial motor generator in the arch and a commercial transformer in the box.

4.5.2 Target variable. The measure of vulnerability was defined as the probability, from 0.0 to 1.0, that the contents of the buried box would have ceased to function after the scenario described.

4.5.3 Subject Matter Experts (SMEs). The SMEs in this study have been described above. They were also the subjects of the study.

4.5.4 Comparison cases. For each target case, three appropriate actual test cases were used. These were chosen from the HEST (Kiger, Slawson, & Hyde, 1984), Kachina (Betz, Smith, Walhood, Edwards, & Verner, 1983), and Dynamic Shear (Slawson, 1984) test series, for which photographs and relevant data were available from AFWL. Two or three photographs were available for each case, showing post-blast condition.

Table 2
BACKGROUND OF SUBJECTS

	PILOT STUDY	PHASE I	PHASE II
<u>Number</u>	4	15	12
<u>Employ</u>			
Government		8	
Industry		7	12
University	4		
<u>Years of Experience</u>			
20 or more	3	6	5
11 - 19		7	4
10 or less	1	2	3
<u>Area of Expertise</u>			
Structures	4	10	9
Soil		2	
Blast		1	
Contents			1
Instrumentation		1	
Soil/Structures		1	1
Contents/Structures			1

4.5.5 Causal factors. The parameters judged to be relevant to the differences found from case to case were classified as follows:

BLAST: Yield, Height of burst, Overpressure
 GEOLOGY: % Sand, % Clay, % Water, % Rock
 STRUCTURE: Roof shape, Length, Width, Height, Radius,
 Concrete strength, Wall thickness, Burial depth,
 Scale, % Steel
 CONTENTS: Equipment, Restraint

4.5.6 Scenario. For each target case a sheet was prepared with a drawing of the geometry of the structure and a listing of particulars for each of the causal factors. If a subject requested more specific data, and put forth an assumption, the assumption was confirmed. If direct requests for data were made, the response was (1) "the same as in the comparison (or target) case" or (2) "those data are missing."

4.6 FINAL CBP STRATEGY

The pilot study interviews had revealed that experts considered the effect of factors in combination, rather than singly, and that a structure to capture or support their expertise could not force a quantification of factor-by-factor effect. It was therefore decided to use a simpler CBP strategy to accommodate that finding.

In the final twelve interviews, subjects were asked, as before, to judge the probability of content failure in each of three comparison cases. They were then asked to place the target case on a continuum constructed of two of these three judgments. That is, they had to make only a relative prediction for the target case, based on a global rather than analytic assessment of the differences between it and the comparison cases. When the target case had been positioned relative to the two comparison cases, the SME then assigned it a vulnerability probability in keeping with its place among them.

Following this assignment of a target value, SMEs were asked to identify those factors which they felt contributed to the difference in vulnerability between the target case and each of the three comparison cases. They rated each factor as to whether it had contributed little, much, or not at all to the difference, and whether its contribution had been to raise or lower the vulnerability rating. The experts were able to make these judgments with little difficulty. The resulting data allowed us to rate factors according to the degree to which experts judged they affected vulnerability, and to identify patterns of factors important to different groups of experts.

5.0 RESULTS

Comparison Based Prediction (CBP) produces two distinct types of data. The first type is numerical data associated with the predictions on the target case and comparison cases. The second type is subjective, qualitative data.

Reported first are results from Phase I and II interviews. Tables 3 and 4 report the mean prediction scores for each group and the associated standard deviations. Prediction scores consisted of the probability, from 0 to 1, that the specified contents of the given structure would have ceased to function immediately following the blast described. These scores were summed and divided by the total to obtain a mean prediction score for each target case for each group of subjects.

Thus in Table 3, the eight subjects who were led through the CBP process on Target Case A2 in Phase I of the study produced a mean prediction score on that case before the structured interview of .58 with a standard deviation of .25. After having applied the CBP process, the same group produced an average vulnerability prediction of .47 with a standard deviation for those judgments of .29. Since the validity of predictions was unknown and could not be tested, no statistical analysis of differences in mean scores was undertaken. Tests of changes in variability were made to determine the effectiveness of CBP in increasing reliability of estimated vulnerability. No significant change was found in these eight.

Table 3
PHASE I RESULTS: PREDICTION SCORES¹

	BEFORE CBP		AFTER CBP	
	Mean	S.D.	Mean	S.D.
² Target Case A2				
Group using CBP (n=8)	.58	.25	.47	.29
A3, non-CBP, same group (n=8)	.46	.28	.57	.35
Target Case A3				
Group using CBP (n=8)	.48	.27	.80	.18 **
A2, non-CBP, same group (n=8)	.58	.25	.56	.29
Target Cases A2 & A3 combined				
Group using CBP (n=16)	.53	.26	.63	.29
A3 & A2, non-CBP, same group (n=15)	.52	.28	.57	.31

¹Prediction scores are estimates of the probability that the specified contents of the structure will have ceased to function immediately following the blast.

²Target cases, detailed in Appendix A, are scenarios of blast, environment, structure, and content for which predictions of contents failure (vulnerability) were made.

**difference in standard deviations is significant at the .001 level.

Table 4
PHASE II RESULTS: PREDICTION SCORES¹

	BEFORE CBP		AFTER CBP	
	Mean	S.D.	Mean	S.D.
² Target Case A2				
Group using CBP (n=4)	.56	.35	.66	.23
A3, non-CBP, same group (n=4)	.69	.34	.69	.34
Target Case A3				
Group using CBP (n=7)	.68	.28	.80	.15 *
A2, non-CBP, same group (n=7)	.41	.31	.56	.32
Target Cases A2 & A3 combined				
Group using CBP (n=11)	.64	.30	.75	.19 **
A3 & A2, non-CBP, same group (n=11)	.51	.33	.60	.32

¹Predictions scores are estimates of the probability that the specified contents of the structure will have ceased to function immediately following the blast.

²Target cases, detailed in Appendix A, are scenarios of blast, environment, structure, and content for which predictions of contents failure (vulnerability) were made.

*difference in standard deviations is significant at the .05 level.
**difference in standard deviations is significant at the .001 level.

The Target Cases are included in Appendix A. They each consist of descriptive data for a structure, an associated blast, and contents placed within the structure. The prediction scores generated by each SME are detailed in Appendix B.

Tables 5 and 6 show the influence weighting each SME gave to the different causal factors he judged important. Following these data in Section 5.4 is a summary of the more subjective results. These include both comments from SMEs about the weight they gave different causal factors, and interviewer observations of SME judgment process.

The Phase II and III interview guide, and case materials, are in Appendix A. Complete tables of Phase II and III data are in Appendix B. These include predictions by each SME on each case, including the photographed comparison cases. A description of the Pilot Study CBP design, and the data it yielded, are in Appendix C.

5.1 PHASE I NUMERICAL RESULTS

A total of sixteen subjects, ten from government and six from industry, were asked to predict the vulnerability of contents in eight buried structures to nuclear weapons effects. The prediction was to be between 0.0 and 1.0, with 1.0 being certain failure of the contents.

Six subjects used CBP for Target Case A2, another six for Target Case A3, and two for both A2 and A3. (Target Cases are presented in Appendix A.) The non-CBP case was then used as their control; that is, if CBP was applied to case A3, then A2 was the control case, and vice versa. Table 3 shows the mean predictions for those cases.

For the group using CBP with Case A2, only minimal differences were obtained between pre- and post- predictions. The standard deviations of their prediction scores for both target and control (A3) cases showed little change.

For the group using CBP with Case A3, a significant difference in pre- vs. post-CBP variability was found for the target A3. The standard deviations for the target predictions were reduced from pre- to post-CBP by one-third. This is significant; the probability for such a difference to occur by chance is less than .001 ($p < .001$).

5.2 PHASE II NUMERICAL RESULTS

Phase II interviews included twelve SMEs from industry, eleven of whom provided usable data. One subject declined to follow the procedure. Each subject was asked to predict the content vulnerability for two hypothetical target cases and two photographed comparison cases. Four subjects had A2 as the experimental case and A3 as the control case, while seven subjects had the opposite. Table 4 shows the mean predictions for these SMEs.

For the group with A2 as the target, there was a drop in the standard deviation for case A2 (.35 to .23) and not case A3 (.34 to .34). For the group using A3 as the target, there was a significant ($p < .05$) decrease in standard deviation. For all eleven subjects combined, the standard deviation was significantly lower in the pre- vs. post-CBP case ($p < .001$) while showing only minimal change in the control case.

5.3 INFLUENCE OF SYSTEM PARAMETERS

The area of expertise for each SME was recorded as well as the influence of key variables on his prediction process. Tables 5 and 6 summarize the Phase I and II SMEs' expertise and the variables they considered to be important. The values for the key variables range between (+)4 and (-)4. A weighting of (+)4 indicates that the variable highly influenced the SME to judge the target case more vulnerable than the comparison case. A weighting of (-)4 indicates that the variable highly influenced the SME to judge the target case less vulnerable than the comparison case. Blank cells indicate the SME did not use that factor.

In general, the SMEs felt only two to three factors were different enough between cases to produce a change in system response, this despite the fact that almost all parameters differed among the various cases. The average number of factors believed to be significant was equivalent for both A2 and A3 with, respectively, means of 2.61 and 2.57 different factors considered by SMEs in making the prediction for each case.

5.4 SUBJECTIVE DATA

The CBP method provides not only predictions but also insights into an SME's method for problem solving or information processing. The insights are reflected in opinions or questions from the SME and observations made by the interviewers. Eight distinct areas of subjective data can be drawn from SME and interviewer comments.

5.4.1 Initial SME Assumptions

The problem solving process was significantly affected by the initial assumptions held by the SME. His initial review of the target cases, and the associated initial predictions, were often biased by the structural responses he assumed. When initial predictions conflicted with evidence in the photographed comparison cases, the SME then recognized his erroneous assumption for the initial prediction.

Among these assumptions was that of structural integrity. SMEs assumed the target structure would survive unbreeched. They then saw the photographed comparison cases, revealing structural damage after similar blasts. This invalidation of the initial prediction assumption seemed to produce large changes in pre- vs. post-CBP predictions.

5.4.2 SME Data Aggregation

There was a marked difference between SMEs in the handling and combining of system variables. Certain SMEs preferred to look at each variable individually, while others combined the parameters into one variable to reduce interaction effects. For example, in the evaluation of a buried arch response, some SMEs considered radius and wall thickness individually while others combined them into one variable.

Table 5
CAUSAL FACTOR EMPHASIS
CASE A2

TARGET CASE: A2

	AREA	COMPARISON:	STRUCTURE										COL	BLAST	CO	W
	EMPH OF	CASE	LENTH/width/space	ICONC.	SPAN/DEPTH/REACTS	WALL	REGS-ADJUST/LENTH	BURIAL/TYP	IMPULSE	YIELD/LOADING	OVER-	% OF FACTORS				
	EXPERTISE	NUMBER	WIDTH	STEEL/STRENGTH/INCNS/width	STEEL/STRENGTH/INCNS/width	INCNS/width	INCNS/width	INCNS/width	INCNS/width	INCNS/width	INCNS/width	PREDICTED	TOTAL			
5	STRUCT	+	+	-0.5	+	+	+	-0.5	+	0.5	+	+	-0.5	+	+	4
5	STRUCT	+	+	+	+	+	+	1.5	+	0.5	+	+	1.5	+	+	2
6	STRUCT	+	+	+	+	+	+	-0.5	+	1.5	+	+	-0.5	+	+	2
6	STRUCT	+	+	+	+	+	+	-0.5	+	1.5	+	+	-0.5	+	+	2
7	STRUCT	+	+	+	+	+	+	-0.5	+	1.5	+	+	-0.5	+	+	2
7	STRUCT	+	+	+	+	+	+	-1	+	-0.5	+	+	-1	+	+	2
8	STRUCT	+	+	+	+	+	+	1.5	+	1.5	+	+	1.5	+	+	3
8	STRUCT	+	+	+	+	+	+	-4	+	0	+	+	-4	+	+	2
9	BLAST	+	+	-1	-1	+	+	-4	+	0	+	+	-4	+	+	2
9	BLAST	+	+	+	+	+	+	2	+	0	+	+	2	+	+	2
10	EMPT	+	+	+	+	+	+	1	+	0.5	+	+	0.5	+	+	1
10	EMPT	+	+	+	+	+	+	0	+	0	+	+	0	+	+	1
10	EMPT	+	+	+	+	+	+	-1	+	0.5	+	+	0.5	+	+	4
11	STRUCT	+	+	+	+	+	+	-1.5	+	0	+	+	-1.5	+	+	3
11	STRUCT	+	+	+	+	+	+	-4	+	0	+	+	-4	+	+	2
12	STRUCT	+	+	+	+	+	+	3	+	0	+	+	0	+	+	2
12	STRUCT	+	+	+	+	+	+	0	+	0	+	+	0	+	+	2
13	STRUCT	+	+	+	+	+	+	0.5	+	0	+	+	0.5	+	+	1
13	STRUCT	+	+	+	+	+	+	-0.5	+	0	+	+	-0.5	+	+	1
14	STRUCT	+	+	+	+	+	+	0.5	+	0	+	+	0.5	+	+	2
14	STRUCT	+	+	+	+	+	+	-0.5	+	0	+	+	-0.5	+	+	2
15	STRUCT	+	+	+	+	+	+	-1.5	+	0	+	+	-1.5	+	+	3
15	STRUCT	+	+	+	+	+	+	-4	+	0	+	+	-4	+	+	2
16	STRUCT	+	+	+	+	+	+	-1.5	+	0	+	+	-1.5	+	+	3
16	STRUCT	+	+	+	+	+	+	-4	+	0	+	+	-4	+	+	2
17	CONTENTS	+	+	+	+	+	+	0.5	+	0	+	+	0.5	+	+	2
17	CONTENTS	+	+	+	+	+	+	-0.5	+	0	+	+	-0.5	+	+	2
18	STRUCT	+	+	+	+	+	+	-0.5	+	0	+	+	-0.5	+	+	2
18	STRUCT	+	+	+	+	+	+	-4	+	0	+	+	-4	+	+	2
19	STRUCT	+	+	+	+	+	+	-1.5	+	0	+	+	-1.5	+	+	2
19	STRUCT	+	+	+	+	+	+	-1.5	+	0	+	+	-1.5	+	+	2
20	STRUCT	+	+	-1.5	+	+	+	-1.5	+	-1.5	+	+	-1.5	+	+	3
20	STRUCT	+	+	-1.5	+	+	+	-1.5	+	-1.5	+	+	-1.5	+	+	3
21	CONTENTS	0	1	1	1	1	2	2	2	2	2	2	2	2	2	2

Table 6
CAUSAL FACTOR EMPHASIS
CASE A3

TARGET CASE: A3

5.4.3 Targets as Comparisons

As indicated previously, subjects used all the original hypothetical cases as comparison cases. Subjects referred back to previous cases when making initial predictions.

Even when given only two preliminary cases, SMEs referred back to the first of these cases to make the predictions on the second. The SMEs were therefore assuming the validity of their initial prediction, and ensuring all subsequent predictions were consistent with the initial one, thus constraining subsequent initial predictions.

This SME response characteristic indicates that SMEs do process information by analogical reasoning in some form. This analogical processing, using comparisons, occurred prior to any mention of the project purpose or theories.

5.4.4 Failure Modes

One of the initial assumptions made by SMEs was the mode of failure for the contents. In general, the contents were judged to fail due to either (1) acceleration/velocity effects, or (2) impact failure (roof falling on contents). At least one SME changed his judgment of the most likely mode of failure for a target case because of the evidence on photographed comparison cases. He had initially based probability of failure on acceleration, from the first comparison case effects. He changed the probability of failure, after viewing a second comparison case, and also changed the basis for that judgment to impact failure (roof collapse), without considering acceleration effects. And yet, acceleration was continuing to increase as the environmental stress increased. The SME did not combine the two key probability failure curves.

Figure 1 illustrates the potential impact of this strategy on the prediction. Two hypothetical curves are drawn, for the relation of structural failure and of acceleration damage to the probability of contents failure. A third curve indicates the possible relation between a combination of the two types of damage and the probability of contents failure. This combined curve indicates a higher probability of contents damage than either of the single-factor curves, at every point.

In no predictions did any SME combine failure mode probabilities. The prediction was based on one curve only, even though the other failure mode potential was present. This indicates to us that judgment of vulnerability is likely to be generally underestimated.

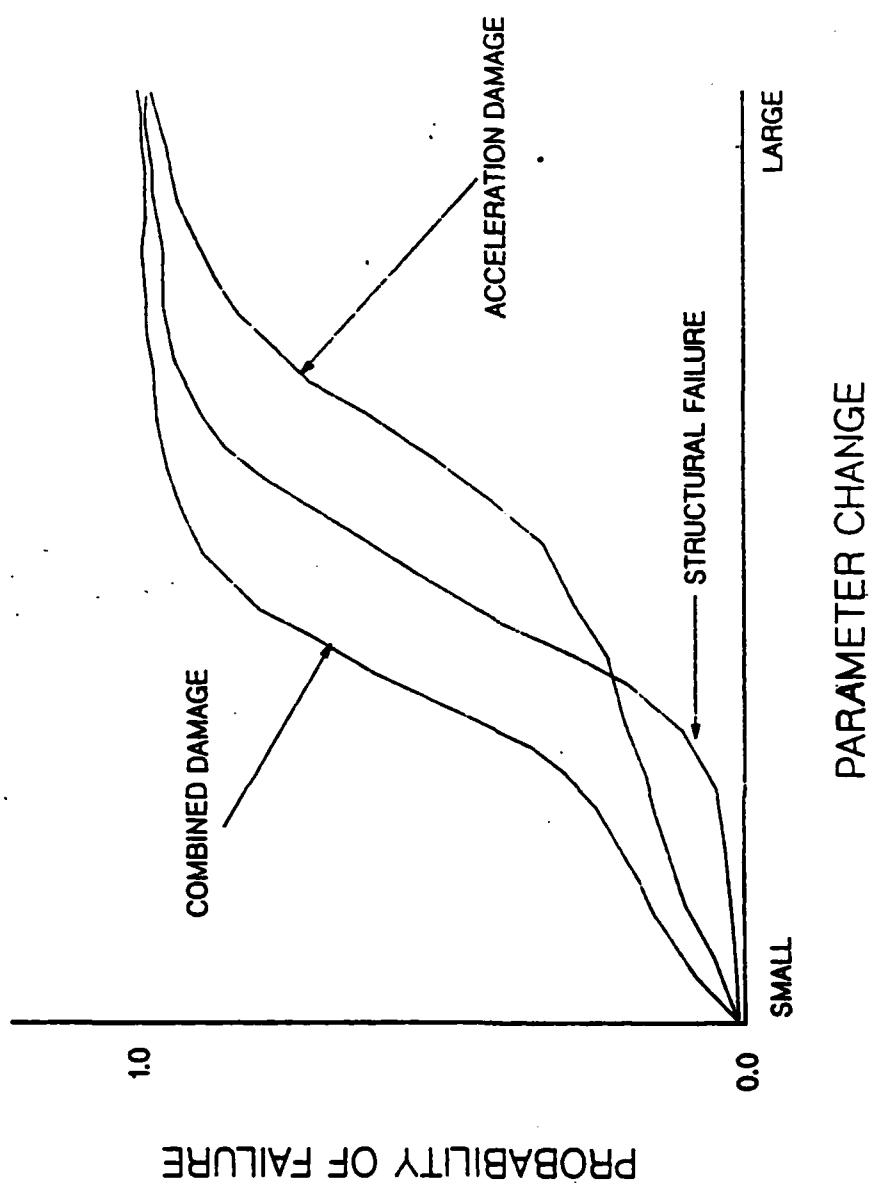


Figure 1
SEPARATE AND COMBINED FAILURE MODE ESTIMATES

5.4.5 Construction Quality

Several SMEs indicated that differences between comparison case responses were due to construction quality and not design parameters. For example, two similar structures showed radically different responses to slightly (20%) different stressors. Several SMEs believed that neither the design nor environmental differences were significant, and the different structural response was due to quality of construction. Other SMEs argued later that construction quality was closely monitored. However, SMEs could not rule out quality differences as a major source of variance in structural response.

5.4.6 Manipulation of Data Outside Expertise

Since the predictions encompassed the entire S/V problem, each SME was confronted with a problem outside his area of expertise. Two typical SME responses were observed when faced with "unfamiliar" data.

In one typical response, SMEs erroneously or simplistically interpreted data on variables not in their area of experience. For example, one SME felt a reported soil water content of 12% was wet, while another felt 7% was very dry. It is understood that water percentage, while important, should not be considered in isolation; however, several SMEs did not ask for the additional required data (e.g. soil fracture characteristics) or even state that certain assumptions had to be made. In another area, most structurally oriented SMEs felt a transformer was less vulnerable than a motor generator (M/G) because the M/G had more moving parts. A contents-expert indicated that, while an M/G does have more moving parts, the transformer was more sensitive to velocity/acceleration effects, because of the position-sensitive nature of the transformer core.

In a second typical response, SMEs felt parameter differences outside their area were more important than did the SMEs who were knowledgeable in that area. SMEs not experienced in soil/structures considered slight soil changes to be important, when experts in that area felt the same changes to be of no significance. Because "area of expertise" was defined by the SME, it is not clear how much or how little experience SMEs had outside their stated area.

5.4.7 Missing Data

In order to have manageable cases, the targets and comparisons were simplified. All possible comparison and target variable data were not provided. Differences were observed among SMEs in what they felt was important data that had not been included.

Again, area of expertise strongly influenced the questions regarding missing data. Structural experts wanted information on structural parameters. The characteristics of the rebar were frequently requested. Content experts felt acceleration information and equipment response spectra to be critical to prediction, and asked for these data. Of note is that content experts did not ask for structural data and vice versa.

5.4.8 Non-Linearity

One of the primary obstacles to the prediction process was the non-linearity of the problems under consideration. That is, the combinations of factors, and changes in important factors from case to case, did not yield straight line effects. The ability to assume linearity, or a straight line effect of changes, facilitates the prediction process.

The pilot study required SMEs to look at the relationship of the target case to three comparison cases at once. Phases I and II allowed one of these cases to be discarded and the SME to assume a linear relationship between the remaining comparison cases, at least on the 0.0 to 1.0 scale of vulnerability prediction. This made it easier to position a predicted vulnerability for the target case.

Figure 2 shows a hypothetical response of a structure to a given parameter change, with an obvious non-linear relationship present. However, if only two points on the curve are shown, as essentially was done in Phases I and II, then a person can infer or assume a linear relationship. It is obvious that the assumption of linearity can significantly affect the prediction process.

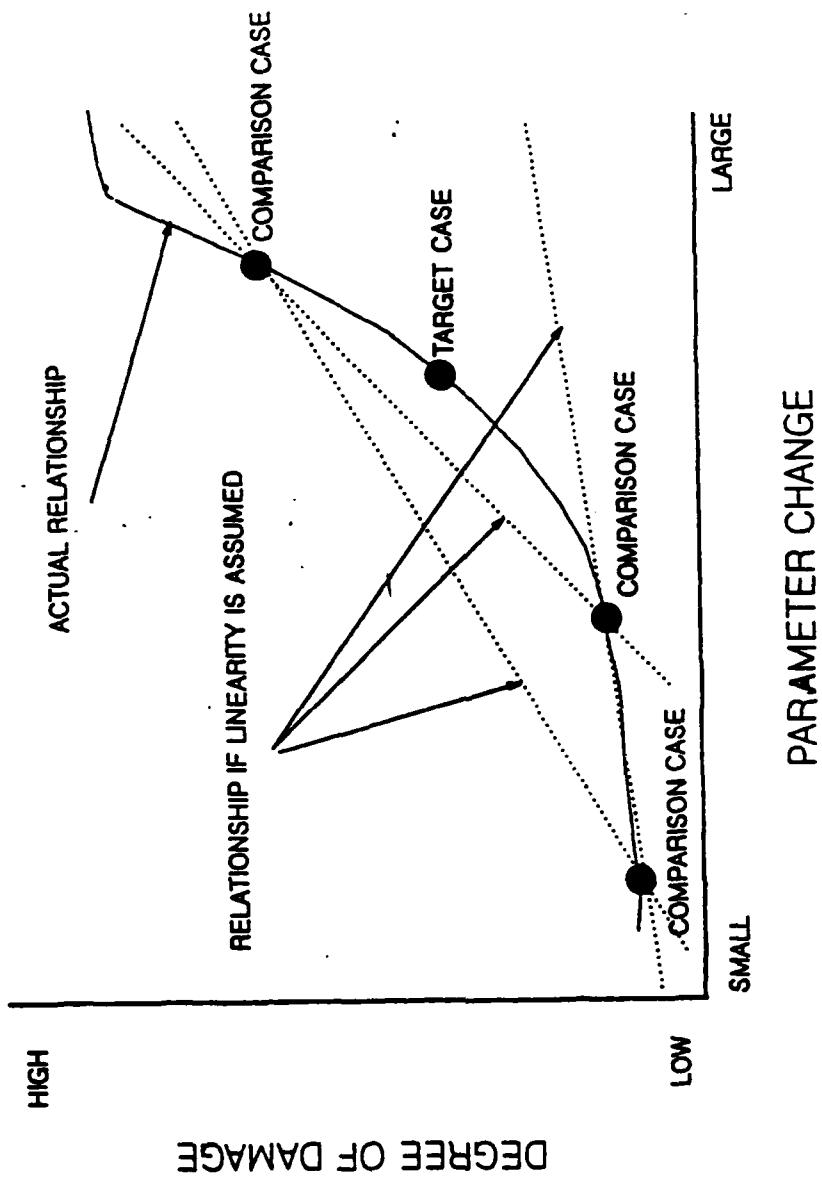


Figure 2

CONTRAST OF LINEAR vs NON-LINEAR ASSUMPTIONS ON ESTIMATED
DEGREE OF DAMAGE

6.0 DISCUSSION

The major finding was that Comparison-Based Prediction was a feasible method for generating S/V estimates. Once the procedure was developed for use with S/V problems, twenty-seven subjects were interviewed, and of these, twenty-five were able to use the method to organize their judgments within its guidelines and, using its required elements, generate a prediction. Moreover, the subjects reported that they felt comfortable using the method, and gave support to our findings in earlier research (Klein & Gordon, 1984) that subjects tended to prefer their CBP-structured judgments to their unstructured judgments. This effect was more striking for the last twelve subjects that were interviewed, after we had altered the design. The one SME from this group who was not able to generate a prediction was uncomfortable with the way the exercise was conducted, not with the CBP method. The sample problems were too simplistic for this SME to use his experience with, and some data that he felt were critical were not available. (One interview in Phase I had been interrupted and that SME's prediction was not completed.)

Since we did not have the opportunity to generate predictions about actual S/V experiments, we were unable to study the validity of the method. However, we did have a chance to measure its reliability. The effect of the CBP method was striking in reducing the variability of SME judgments. The standard deviations were significantly reduced using the CBP method as opposed to the initial unstructured judgments. Previous research (Klein & Gordon, 1984) had found that the standard deviations were reduced by 25% using CBP, and in the present study, the standard deviations were reduced by over 30% for each of the two target cases in Phase II, a result that was significant at the .001 level for the two cases combined.

It is also interesting to trace the evolution of the procedure as we conducted our interviews. The initial pilot interviews asked SMEs to provide specific estimates for causal factors, and we found that they were unable to generate such values. (This may have implications for knowledge engineering approaches that depend on such data.) Therefore, we shifted our paradigm to one where subjects made over-all comparisons between a target case and several analogue cases. These comparison cases served to bound the prediction range. This method was effective, but our design had a weakness in that we presented several comparison cases for initial estimates, planning to contrast these with estimates obtained after CBP. The problem was that the SMEs used the initial estimates to make direct comparisons, thus introducing CBP into the "unstructured" judgments. This reflects subjects' normal tendency to rely on relative and comparative judgments where possible. This is the process that we have sought to capture using CBP: the normal use of analogical reasoning. We are trying to formalize this use, to make it explicit and more objective. It was only after we changed our design to prevent the SMEs from making direct comparisons during the initial estimates that we were able to show how the formal CBP method was consistently reducing the variability of judgments. But even in Phase I, the use of CBP reduced variability significantly for one of the two target cases.

Our data also allowed us to study the way that the SMEs generated their predictions. This met the revised objective of using CBP as a knowledge elicitation tool, to be used to examine the bases for subjects'

judgments. In general, we confirmed that the SMEs did use analogical reasoning. Support for this is found in the reduced standard deviations obtained through CBP, and the observation in the Phase I design that the SMEs depended heavily on the comparison cases that we had made available during the initial estimates.

More important, we were able to use CBP to trace the causal factors used by each SME in generating judgments. These data were shown in Tables 5 and 6. Thus, SME 13, working with Target Case A2, relied on differences in geometry to adjust comparison cases B5 and B6 (photographs of these boxes, and relevant data, are reproduced in Appendix A); in addition, SME 13 relied on differences in overpressure for comparison case 6. We have not only a record of the causal factors used, but the extent of their influence. The magnitude of influence was rated on an ordinal scale, and the SMEs were able to generate these estimates, whereas for the pilot study the SMEs could not generate interval-scale estimates. This type of data can be valuable for the development of analytical systems based on exemplars to capture the available expertise.

It is worth noting that Tables 5 and 6 show the number of causal factors used by the SMEs. Generally, it is between two and four. In only four cases out of fifty were more than four causal factors employed. This reflects the information processing limits of SMEs to handle several dimensions simultaneously. In addition, there were seven cases where only one causal factor was adjusted. For example, SME 30, working with Target Case A3, focused entirely on blast impulse and disregarded all other influences. No other SME working with the same target and comparison cases was this simplistic.

The specific causal factors used are largely a function of the target and comparison cases used. For example, some SMEs listed a key variable but felt it did not differentiate the target and comparison cases used, and therefore did not apply it. The scenarios prepared for this research were simplistic in many ways, and certainly would not be representative of the S/V domain. The development of an exemplar-driven analytical approach would require a fuller and richer set of cases to serve as exemplars, although the form of the SME judgments could follow the CBP approach.

The CBP method was also useful in identifying the SMEs' assumptions. Thus, some SMEs had examined the scenario data and concluded that the force of the blast was not sufficient to create roof collapse in a comparison case, or in other cases was sufficient, only to see from the photograph that they had been in error. Most SMEs were not expert regarding contents, and assumed that the generator would be more vulnerable than the transformer, attaching greater influence to moving parts than to parts that were easily damaged due to movement. The influence of water content in soil was another case of erroneous SME assumptions coming out during the data collection. Other areas were placement of the contents, restraints on the contents, and assumptions about cratering effects and about the quality of the structures.

It is not clear how to compare SMEs who are knowledgeable in an area to SMEs who are not. For instance, a structures expert might use causal factors pertaining to geometry, and so might a subject who does not work in structures but who is concerned about generalizing from an arch to a box.

In some ways, experts and non-experts were remarkably similar in their use of causal factors, and in other ways they were clearly different.

Thus, in forty-two out of fifty cases where there was multiple reliance on the same causal factor (see Tables 5 and 6), the non-expert was in clearly the same range as the expert(s). In only eight was there a difference, and in all of these there had been a difference between the experts themselves. In some of these cases, the non-expert was exactly midway between the experts. In no cases did the non-expert disagree with a set of consistent expert judgments.

On the other hand, there were differences between experts and non-experts. Soil type was given more emphasis by non-experts than by experts, and for blast characteristics both loading and yield were judged to be less important by experts than by non-experts.

This issue raises the question of how we identify SMEs as experts in particular areas. This was done using two criteria. One was acknowledged experience. We generally relied on the content expertise attributed by others, such as AF specialists, and by the SMEs themselves who knew where their skills were. Second, we were able to verify these estimates by looking at the issues that the experts were raising. The true experts, in our judgment, were able to distinguish between causal factors in greater detail, and to recognize the influence of differences. For example, there was general agreement that the motor generator would be more vulnerable than the transformer. However the SME whom we felt to have the greatest expertise in contents disagreed. He felt that the transformer was more sensitive to acceleration effects. At first, this surprised us (especially since he was one of the last SMEs run), but after listening to his description of avenues for component failure it became clear that the other SMEs, knowledgeable in issues such as blast, soil, and structure, were not sensitive to contents and had merely been distinguishing between moving parts vs. no moving parts.

This example about contents is valuable for showing the importance of specific types of expertise. Here it would have been critical to have only the appropriately expert SME make the judgment.

7.0 CONCLUSIONS AND RECOMMENDATIONS

To summarize our evaluation of CBP in this context, it appeared to be very effective for structuring the expert judgments, as shown by the clear reduction in the variability of the estimates. Furthermore, it was useful for articulating the basis of the judgments, that is, for identifying the causal factors that were going into each of the SMEs' predictions.

These accomplishments suggest several applications for the CBP method as applied to S/V predictions.

(a) First, the CBP method can be effective as a means for knowledge elicitation. Properly applied, it can identify and assign preliminary weights to the factors that would go into a prediction model.

(b) Further, it can have value for structuring the judgments that would go into an expert system. If an expert system were developed, a CBP method could be useful for performing the knowledge engineering to elicit specific patterns of judgments for individual SMEs working with a variety of sample cases. This could provide the input for the expert system.

(c) Where predictions may be required, the CBP method is available as a formal means of extrapolating from a data base. CBP uses operational data to generalize to a new situation, in the way extrapolation is usually accomplished. Without CBP this is informal and non-observable. With CBP, the process becomes structured and an audit trail is constructed. Thus, we were able to make explicit some of the assumptions made by the SMEs. For example, we were able to show that SMEs were basing predictions of survivability purely on acceleration effects, but when they were shown evidence of structural failure, they immediately ignored acceleration effects and focused their predictions entirely on the structural failure. It should be possible to devise CBP strategies to overcome such decision biases on the part of the SMEs. The use of CBP for predictions may be especially effective for issues such as opponent missile survivability in a Strategic Defense Initiative environment.

(d) The CBP approach can also have some value for the development of training programs. The approach creates an audit trail of the causal factors used by experts to make predictions. If these data are appropriately organized, new personnel can study these aspects of SME judgment processes. Training scenarios can be developed for which trainees attempt to generate predictions for target and comparison cases, then comparing their use of causal factors to that of the experts.

In order to develop CBP into a more effective tool for Survivability/Vulnerability analysis, a useful next step would be to perform research that bears directly on the predictive validity of the method. Such research could establish a prediction problem, and contrast predictions made in advance using several different methods. Thus, CBP could be used by itself, or in combination with fuzzy set logic. Several SMEs expressed interest in using more analytical tools and references, and we could study ways of combining CBP with these. Results with these strategies could be compared to predictions obtained by asking experts for estimates without relying on CBP or these other formalized approaches.

We feel that adding the CBP structure would improve the accuracy of current predictions. However, it is not enough to improve accuracy, or even reliability. The accuracy must be improved enough to make the procedure worthwhile. We are not speaking simply of statistical significance. We have found in one study (Klein & Gordon, 1984) that under some conditions predictions correlated with the actual results at a high rate (.76 and .84), and in other conditions the correlation was lower (.36 and .46). All of these were statistically significant. However, a correlation of .36 accounts for less than 15% of the variance. Is this sufficient to warrant the effort of performing the CBP procedure? These questions will have to be addressed in advance, but it is critical to examine the relative value of CBP in improving prediction quality. At present, we can conclude only that the method has value for increasing reliability and for making the basis of predictions apparent by constructing an audit trail of the causal factors going into the

prediction. These merits may be important in their own right, but they should not be mistaken for predictive accuracy.

In conclusion, this SBIR Phase I effort demonstrated the feasibility of using CBP to address some of the key issues for its application to Survivability/Vulnerability analysis. Many questions remain, along with opportunities for expanding the use of the CBP method.

The final assessment of success for this effort is an evaluation of the accomplishment of objectives. In the SBIR Phase I Proposal, the Primary Objective was stated as determining the feasibility and value of using CBP for S/V issues. With regard to feasibility this has been successfully accomplished. With regard to the value of the method, this has been accomplished for reliability but not for predictive accuracy. With reference to predictive accuracy, the Secondary Objective in the Phase I proposal was to apply CBP to a current AFWL prediction problem, and to conduct a validation study. Because of the limited time frame of the Phase I effort, it was not possible for AFWL to identify such a problem for us to study. Therefore, this objective could not be accomplished. In its place, by examining the use of causal factors by the SMEs, we were able to show how the CBP method can identify the bases for expert judgments, and be an effective tool for knowledge elicitation in S/V analysis.

REFERENCES

Betz, J. F., Smith, J., Walhood, G., Edwards, J., and Verner, J. (1983). Kachina test series: Eagle dancer. AFWL TR- 83-35, Vol. 1, Final Report, July.

Kiger, S. A., Slawson, T. R., and Hyde, D. W. (1984). Vulnerability of shallow-buried flat-roof structures. Department of the Army Technical Report SL-80-7, Report 6, September.

Klein, G. A. (1982). The use of comparison cases. IEEE 1982 Proceedings of the International Conference on Cybernetics and Society, 88-91.

Klein, G. A. (1986). Validity of analogical predictions. Unpublished manuscript.

Klein, G. A. and Gordon, S. E. (1984). Using comparison-based methods for predicting and designing. Proceedings, Psychology in the DoD Symposium, April.

Klein, G. A. and Weitzenfeld, J. (1982). The use of analogues in Comparability Analysis. Applied Ergonomics, 13, 99-104.

Ross, T. and Wong, F. (1985). Survey of failure mode and damage degree analysis. Study prepared for the Air Force Weapons Laboratory, Kirtland AFB, New Mexico.

Slawson, T. R. (1984). Dynamic shear failure of shallow-buried flat-roofed reinforced concrete structures subjected to blast loading. U.S. Army Engineer Waterways Experiment Station Technical Report SL-84-7, April.

Tetmeyer, D. C. (1976). Comparable item approach to establishing frequency of maintenance and maintenance tasks for a new aircraft. Wright-Patterson AFB, OH, Aeronautical Systems Division, Crew Station, Escape and Human Factors Branch.

Weitzenfeld, J. (1984). "Valid reasoning by analogy: Technological reasoning." Philosophy of Science, 51, 137-149.

Weitzenfeld, J. and Klein, G. A. (1982). Comparison-Based Predictions. Klein Associates Technical Report 82-40-6.

Widenhouse, C. and Romans, W. W. (1977). A forecasting technique for operational reliability (MTBF) and maintenance (MMT/FH). Technical Report No. ASD-TR-77-28, Aeronautical Systems Division, Air Force Systems Command, Wright-Patterson AFB, OH.

APPENDIX A
INTERVIEW MATERIALS

SME DATA SHEET

SME # _____ TYPE _____ DATE _____ CCs # _____

NAME _____ PHONE _____

POSITION _____

BACKGROUND _____

YEARS OF EXPERIENCE _____

NOTES _____

SUMMARY SHEET

SME# _____ CCs# _____ DATE _____ ka _____

1) Here are our two cases: What is your rough estimate of the probability (from 0 to 1.00) that the contents of each have ceased to function after the blast described?

A3 _____ A2 _____

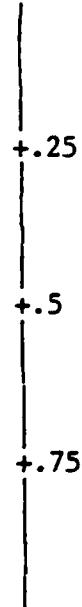
2) We will focus on one of these as our target case. But first, here are two comparison cases on which you will base that prediction. We can show you data and a picture of the results of the blast on the structure.

Please make a rough judgment in each case of the s/v measure for it; that is make a rough judgment of the probability that the contents have ceased to function after the blast described in the structure pictured:

CC# _____ CC# _____

T(B) _____

3) Now, compare the target case, the one which we are trying to predict, with each of these 2 cases. Please tell us whether the probability for the content of the target case ceasing to function is greater, less, or the same for each case, by positioning it in the guide below:



4) Next we would like a few more predictions. Suppose the contents of this structure had been _____, what prediction would you make?

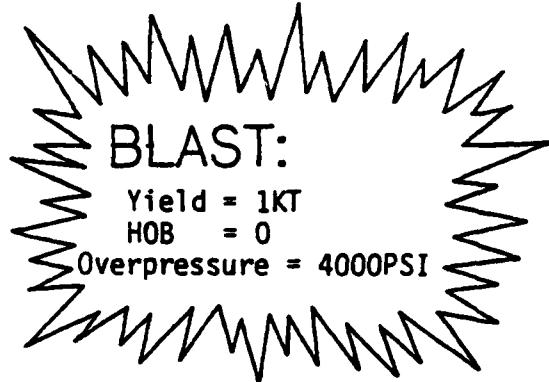
5) Now, please refine your judgment. First, consider the key factors which affect the probability of damage, and the differences in value for each factor between the target case and, one by one, the comparison cases. We will ask you, for each comparison case, to rank evaluate those factors according to how much you think those differences contributed to the difference in outcomes between the two cases.

[done on separate sheets]

6) Here is your original rough estimate of T(A) _____. How do you feel about it, relative to your final prediction(s) using this methodology?

7) Have you any comments on your final estimate, or on this process?

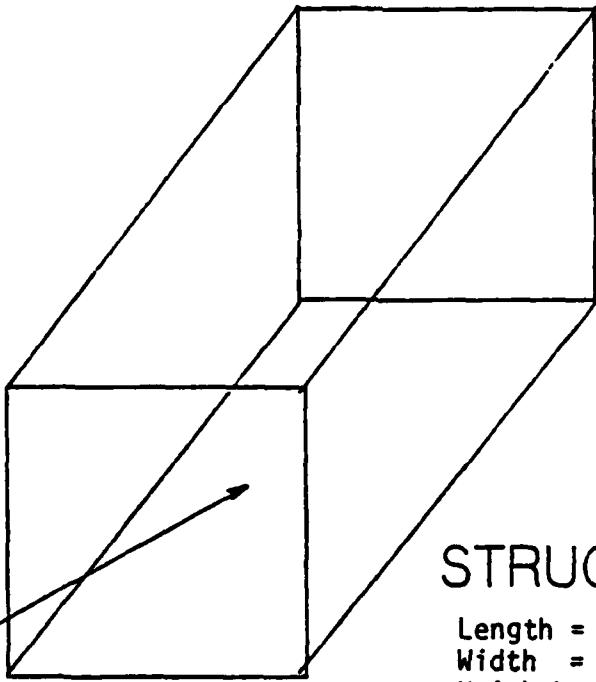
TARGET CASE: A1



GEOLOGY:

SAND = 25%

CLAY = 75%



STRUCTURE: BOX

Length = 15'

Width = 4'

Height = 4'

Wall Thickness = 7"

Burial Depth = 9"

Scale = 1:4

Steel = .9%

Concrete Strength = 5500PSI

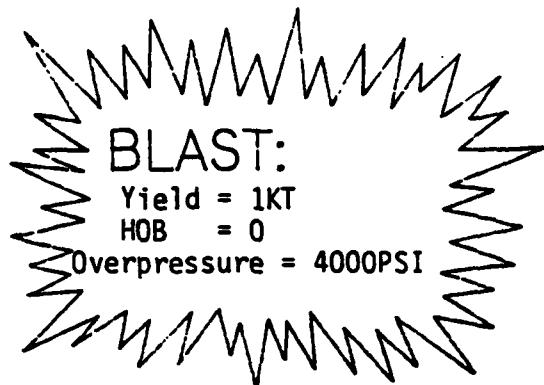
CONTENTS:

IBM PC

Resting on Desk

QUESTION: What is the probability that the contents ceased to function immediately after the blast?

TARGET CASE: A2

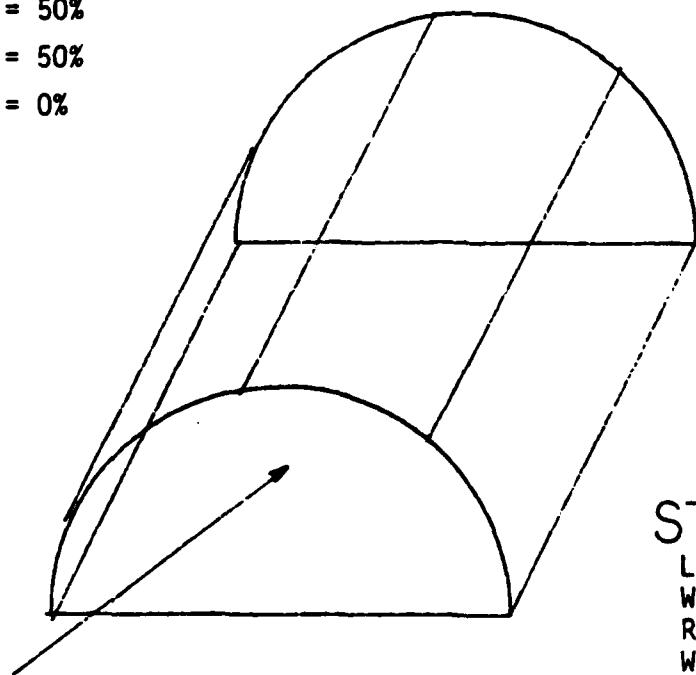


GEOLOGY:

SAND = 50%

CLAY = 50%

ROCK = 0%



STRUCTURE: ARCH

Length = 10'

Width = 4'

Radius = 5'

Wall Thickness = 5"

Burial Depth = 10"

Scale = 1:4

Steel = 1%

Concrete Strength = 5500 PSI

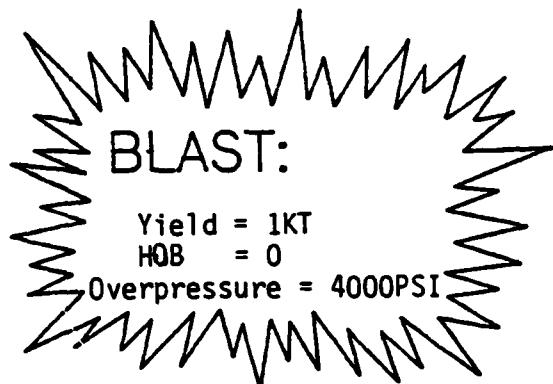
CONTENTS:

Large Motor-Generator
Bolted to Floor

QUESTION: What is the probability that the contents ceased to function immediately after the blast?

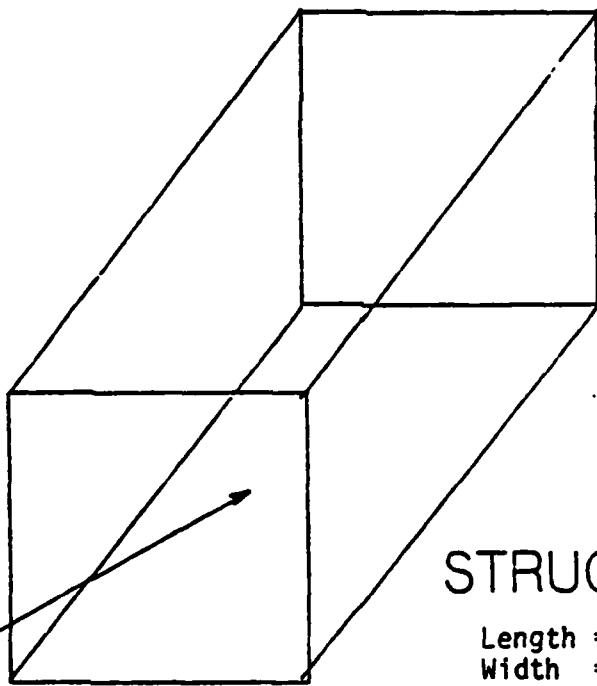
TARGET CASE:

A3



GEOLOGY:

SAND = 75%
CLAY = 25%



STRUCTURE:

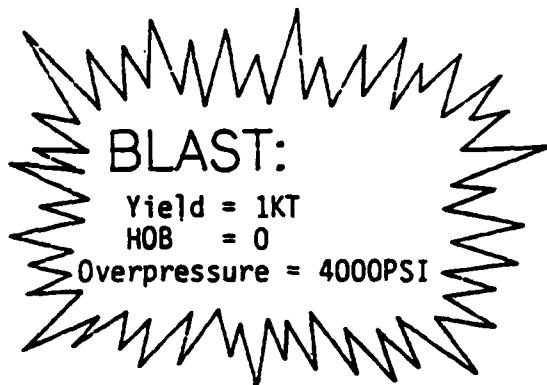
Length = 15'
Width = 4'
Height = 4'
Wall Thickness = 7"
Burial Depth = 9"
Scale = 1:4
Steel = .9%
Concrete Strength = 5500 PSI

CONTENTS:

Large Transformer
Bolted to Floor

QUESTION: What is the probability that the contents ceased to function immediately after the blast?

TARGET CASE: A4

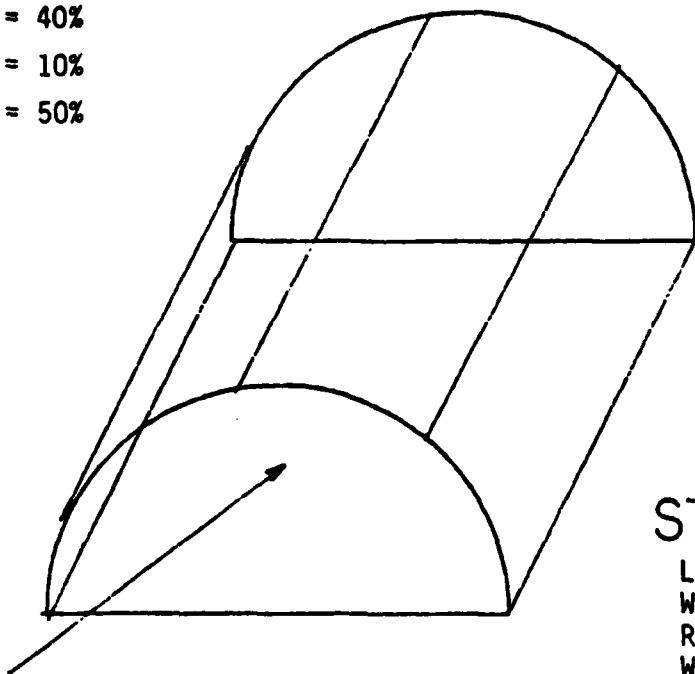


GEOLOGY:

SAND = 40%

CLAY = 10%

ROCK = 50%



STRUCTURE: ARCH

Length = 10'

Width = 4'

Radius = 5'

Wall Thickness = 6"

Burial Depth = 12"

Scale = 1:4

Steel = 1.1%

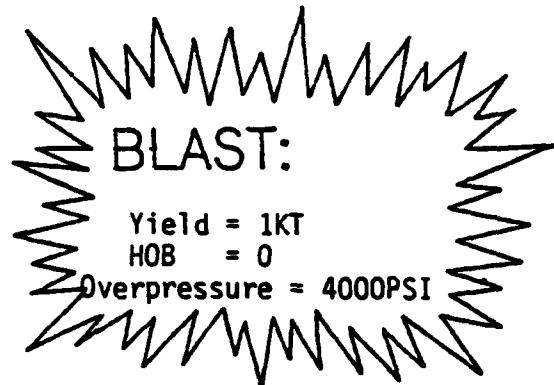
Concrete Strength = 4500 PSI

CONTENTS:

Continuous 460v AC cables
Bolted to Floor

QUESTION: What is the probability that the contents ceased to function immediately after the blast?

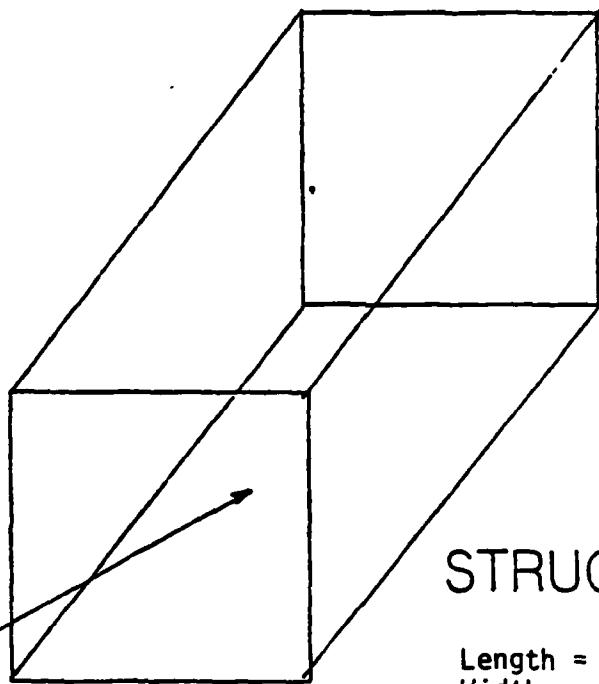
TARGET CASE: A5



GEOLOGY:

SAND = 25%

ROCK = 75%



STRUCTURE:

Length = 15'

Width = 4'

Height = 4'

Wall Thickness = 6"

Burial Depth = 18"

Scale = 1:4

Steel = 1.1%

Concrete Strength = 6500 PSI

CONTENTS:

Person

QUESTION: What is the probability that the contents ceased to function immediately after the blast?

COMPARISON: B4

BLAST

Yield = 1.46
HOB = 0
Overpressure =
3480 PSI

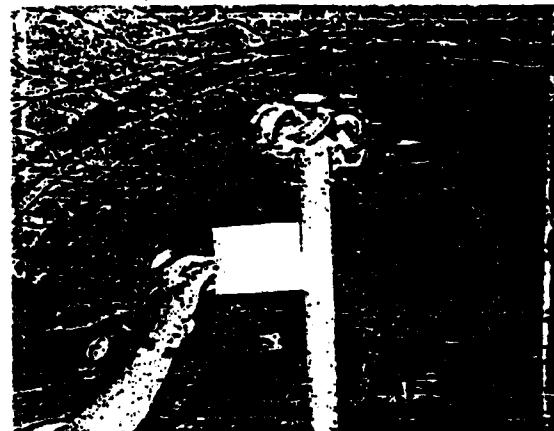
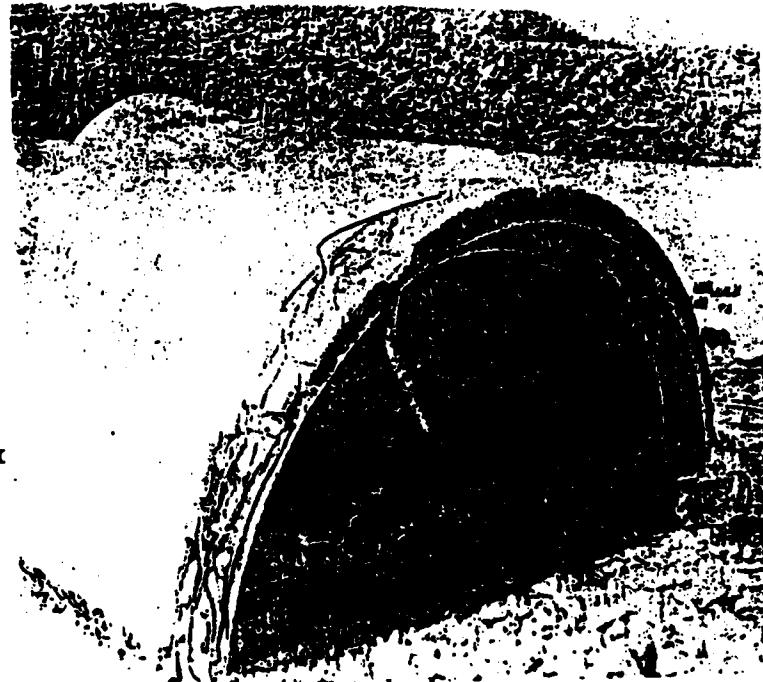
GEOLOGY

Sand-Clay
125 H₂O

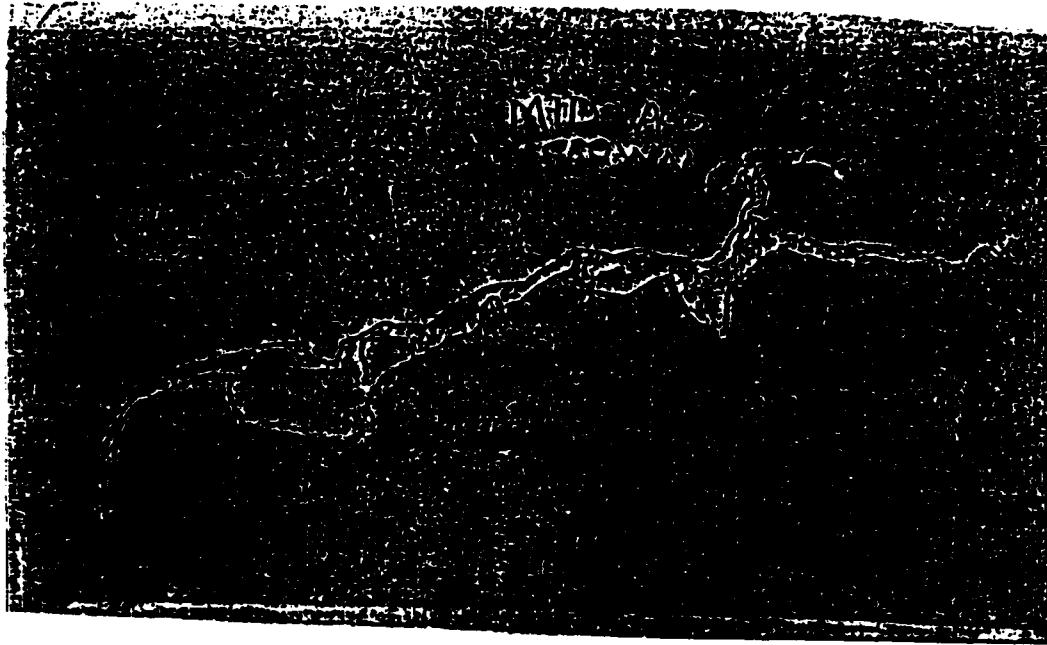
STRUCTURE

Arch
11.15'x7.87'x3.94'
walls = 6.24"
Concrete = 5000 PSI

Rebar
Grade = 120
S Steel = ?
Scale = 1:4
Burial
Depth = 24"



COMPARISON CASE B4 AS PRESENTED TO SMEs (Reduced by one-fifth).
(Adapted from Betz et al, 1983, page 100.)



**COMPARISON CASE B4 page 2 (as presented to SMEs) (Reduced by one-fifth).
(Adapted from Betz et al, 1983, page 99.)**

COMPARISON: B5

BLAST

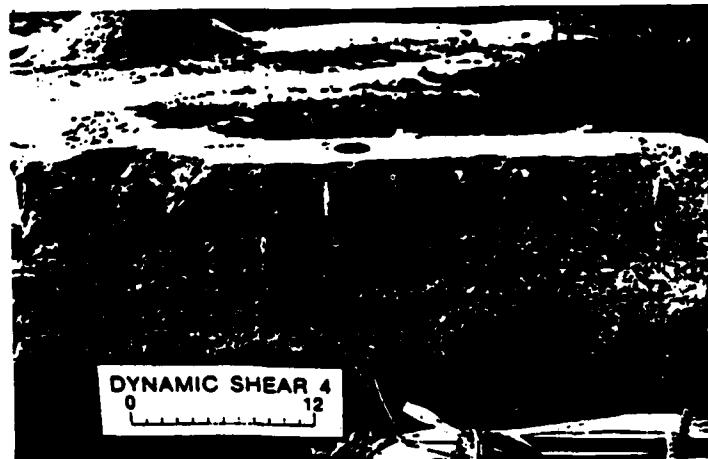
Yield = .84
HOB = 0
Overpressure =
4030

GEOLOGY

Sand
4-7%

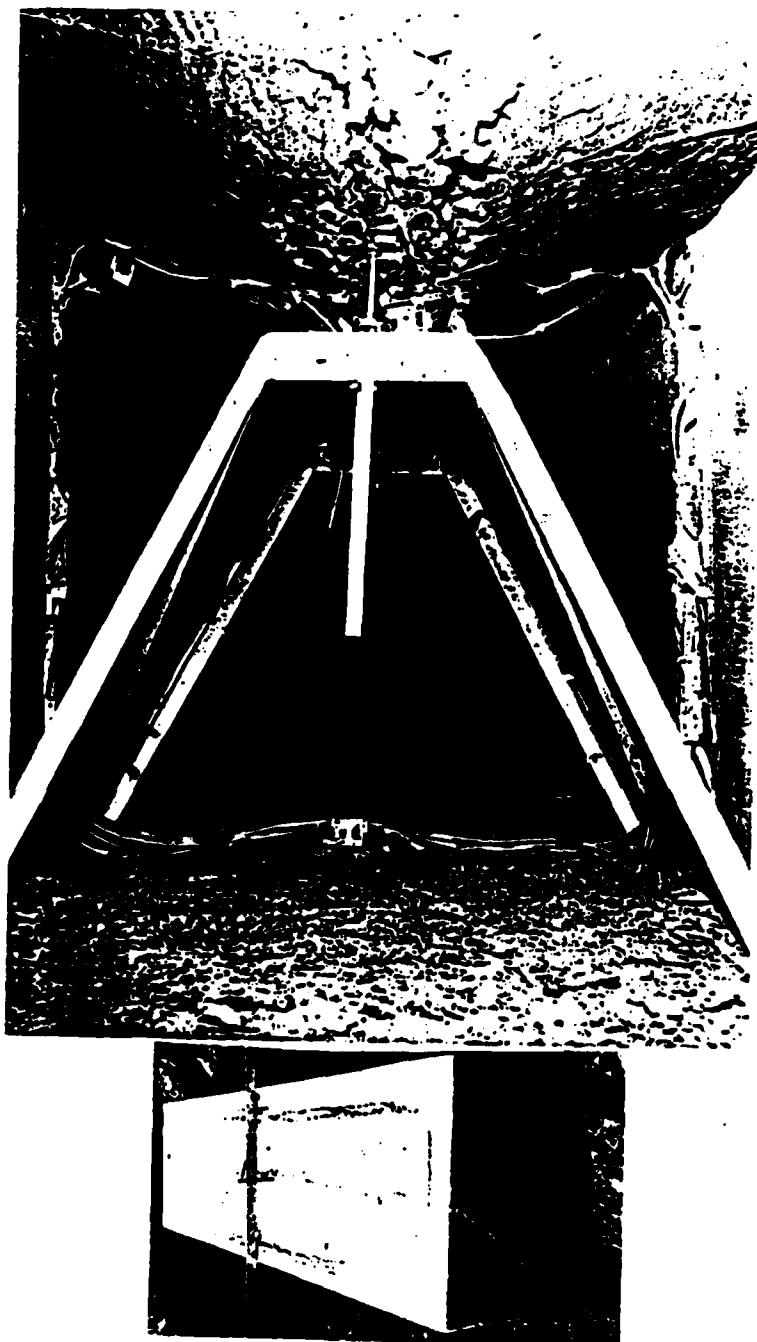
STRUCTURE

Box
4'x4'x4'
walls = 5.6"
Concrete =
6000
Rebar
Grade = 60
S Steel = 1.0
Scale = 1:8
Burial
Depth = 9.6"



COMPARISON CASE B5 A/S PRESENTED TO SMEs (Reduced by one-fifth).
(Also presented as Comparison Case B9)

(Adapted from Slawson, 1984, page 63).



COMPARISON: B6

BLAST

Yield = .08-1.11
HOB = 0
Overpressure =
1600-1900

GEOLOGY

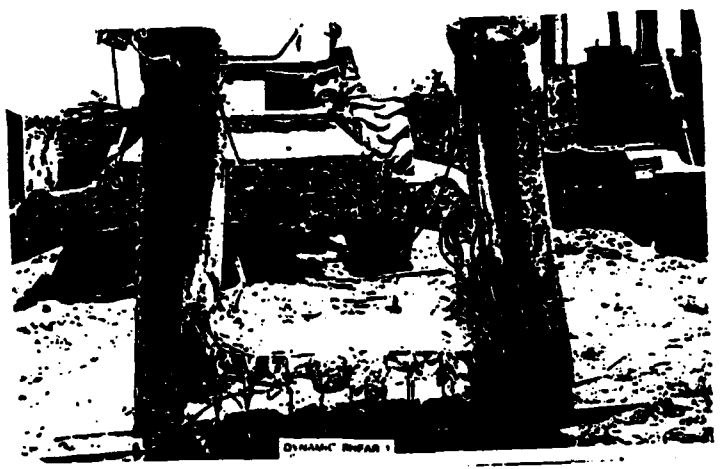
Clay

STRUCTURE

Box
4'x4'x4'
walls = 5.6"
Concrete = ?

Rebar
Grade = ?
S Steel = 1.0
Scale = 1:4
Burial
Depth = 24"

COMPARISON CASE B6 AS PRESENTED TO SMEs (Reduced by one-fifth).
(Adapted from Kiger et al, 1984, pages 92-93.)



COMPARISON: B7

BLAST

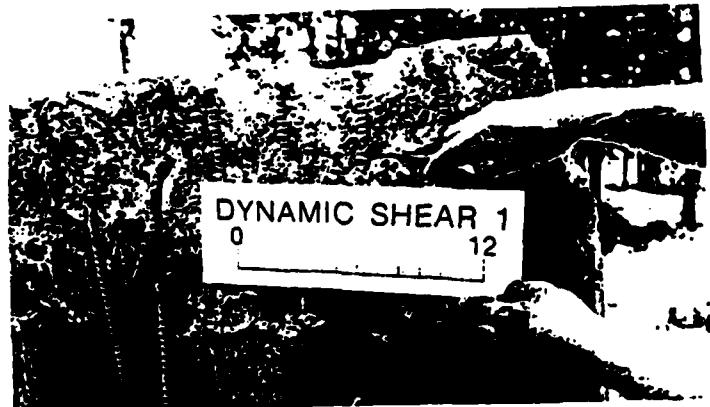
Yield = .99
HOB = 0
Overpressure =
4110 PSI

GEOLOGY

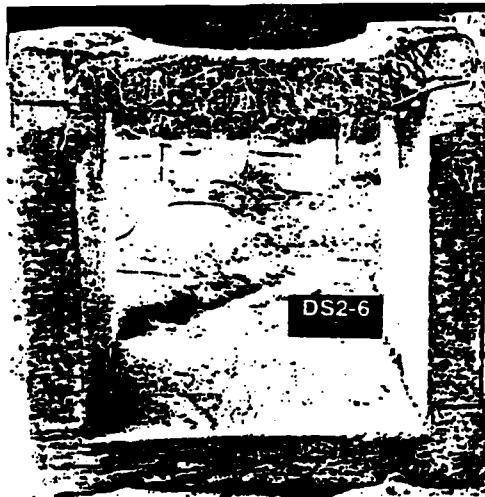
Sand
4-75 H₂O

STRUCTURE

Box
4'x4'x4'
walls = 5.6"
Concrete =
4000
Rebar
Grade = 60
% Steel = 1.0
Scale = 1:4
Burial
Depth = 9.6"



COMPARISON CASE B7 AS PRESENTED TO SMEs (Reduced by one-fifth).
(Adapted from Slawson, 1984, pages 57-58.)



COMPARISON: B8

BLAST

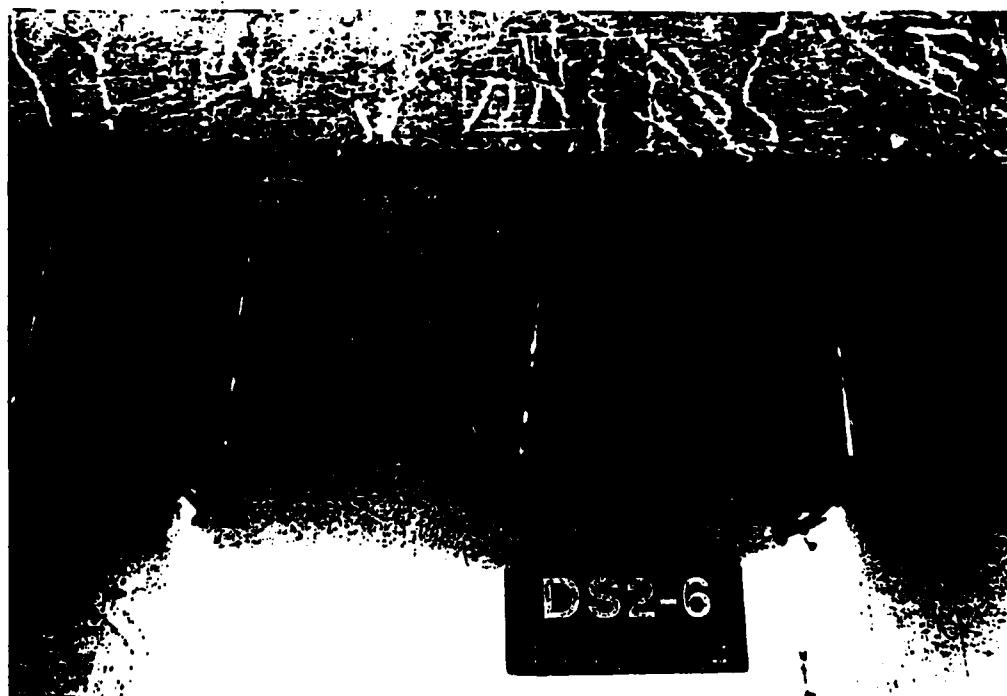
Yield = .8
HOB = 0
Overpressure =
3375 PSI

GEOLOGY

Clay
4-7% H₂O

STRUCTURE

Box	Rebar
3.13'x4'x3.73	Grade = 60
walls = 7.25"	# Steel = 1.2
Concrete = 7000	Scale = 1:4
	Burial
	Depth = 9.0"



COMPARISON CASE B8 AS PRESENTED TO SMEs (Reduced by one-fifth).
(Adapted from Slawson, 1984, pages 75-76.)

APPENDIX B
PREDICTION DATA

Table B:1
PREDICTION SCORES OF PHASE I SMEs

²TARGET CASE: A2

SME	AREA OF EXPERTISE	INITIAL					POST					COMPARISON CASES		
		TA1	TA2	TA3	TA4	TA5	TA1	TA2	TA3	TA4	TA5	TB4	TB5	TB6
9	BLAST	1.00	0.90	0.10	0.01	0.98	1.00	0.95	0.10	0.01	0.98	1.00	0.80	0.90
10	INST	0.80	0.50	0.30	0.01	0.90	0.80	0.35	0.30	0.01	0.90	0.20	0.80	0.20
7	SOIL/STRUC	0.50	0.40	0.30	0.20	0.50	0.50	0.30	0.30	0.20	0.70	0.20	0.90	0.30
6	STRUC	0.90	0.30	0.50	0.20	0.80	0.95	0.10	0.50	0.10	0.80	0.10	0.35	0.10
5	STRUC	0.95	0.30	0.40	0.15	0.90	0.98	0.30	0.45	0.15	0.90	0.25	0.99	0.50
8	STRUC	0.90	0.80	0.30	0.30	0.80	0.99	0.80	0.99	0.30	0.95	0.70	0.99	0.60
13	STRUC	0.98	0.90	0.95	0.95	0.65	0.98	0.60	1.00	0.95	0.95	0.05	0.90	0.50
18	STRUC	0.80	0.50	0.80	0.00	0.10	0.80	0.35	0.92	0.00	0.10	1.00	0.90	1.00

¹Prediction scores are estimates of the probability that the specified contents of the structure will have ceased to function immediately following the blast.

TARGET CASE: AJ

SME	AREA OF EXPERTISE	INITIAL					POST					COMPARISON CASES		
		TA1	TA2	TA3	TA4	TA5	TA1	TA2	TA3	TA4	TA5	TB7	TB8	TB9
18	STRUC	0.80	0.50	0.80	0.00	0.10	0.80	0.35	0.92	0.00	0.10	1.00	0.90	1.00
17	STRUC	1.00	0.90	0.50	0.80	1.00	0.95	0.40	0.40	0.25	0.90	0.95	0.20	0.60
16	STRUC	0.50	0.70	0.30	0.10	0.90	0.90	0.70	0.80	0.10	0.90	1.00	0.50	1.00
15	STRUC	0.80	0.50	0.25	0.25	0.10			0.75			0.95	0.25	0.95
12	SOIL	1.00	0.50	0.50	0.25	1.00	1.00	0.50	0.90	0.25	1.00	0.95	0.25	0.95
14	STRUC	0.70	0.15	0.30	0.01	0.50	1.00	0.15	0.75	0.01	1.00	0.80	0.20	0.80
11	SOIL	0.75	0.50	0.25	0.10	0.95	0.95	0.93	0.85	0.25	0.99	0.75	0.10	0.95
13	STRUC	0.98	0.90	0.95	0.95	0.65	0.98	0.90	1.00	0.95	0.95	1.00	0.05	1.00

²Target cases, detailed in Appendix A, are scenarios of blast, environment, structure, and content for which predictions of contents failure (vulnerability) were made. Comparison cases include photographs of damage to structures after blast.

Table S:2
PREDICTION SCORES OF PHASE II SMEs

2 TARGET CASE: AJ

SME	AREA OF EXPERTISE	INITIAL				POST				COMPARISON CASES			
		TAI	TAE	TAC	TAA	TAI	TAE	TAC	TAA	TAE	TAC	TAA	TBE
20	STRUCT	0.80	0.75			0.90	0.75			0.40	0.90		
21	CONTENTS	0.40	0.20			0.40	0.20			0.30	0.10		
22	STRUCT	0.15	0.95			0.70	0.60			0.40	0.90		
23	STRUCT	0.90	0.80			0.90	0.80			0.70	0.80		

TARGET CASE: AJ

SME	AREA OF EXPERTISE	INITIAL				POST				COMPARISON CASES			
		TAI	TAE	TAC	TAA	TAI	TAE	TAC	TAA	TAE	TAC	TAA	TBE
24	STRUCT	0.10	0.20			0.70	0.95			1.00	0.90		
25	SOIL/STRUCT	0.10	0.50			0.10	0.75			1.00	0.20		
26	CONTENT/STRUCT	0.75	0.85			0.75	0.85			0.70	0.25		
27	STRUCT	0.50	0.90			0.70	0.95			1.00	0.90		
30	STRUCT	0.60	0.90			0.60	0.80			0.95	0.25		
31	STRUCT	0.75	0.90			0.75	0.80			0.99	0.35		
29	STRUCT	0.10	0.50			0.10	0.50			1.00	0.25		

1 Prediction scores are estimates of the probability that the specified contents of the structure will have ceased to function immediately following the blast.

2 Target cases, detailed in Appendix A, are scenarios of blast, environment, structure, and content for which predictions of contents failure (vulnerability) were made. Comparison cases include photographs of damage to structures after blast.

APPENDIX C
INITIAL CBP STRATEGY

The "High Driver" strategy of CBP application was the first choice in this study, and was used in the first four interviews. In this strategy, the expert identifies, or is given, the prime factors on which the test case differs from the comparison case. He/she then attempts to quantify, factor by factor, the effect of these differences on the target value.

Taking the the target measure -- in this case, the vulnerability of contents in the target case -- the SME first examines the known data of the comparison case. In this case, the expert saw photographs of post-blast comparison cases and estimated the vulnerability of a piece of equipment had it been inside the structure. He then adjusted that estimate to fit the target scenario, increasing or decreasing it on the basis of differences between each comparison case and the target case.

For example, viewing the damage to a structure in one comparison case, the expert may judge that the equipment would have a 90% chance of failing to function had it been inside. Comparing that case with the target case, the expert might say that, on the basis of the difference in roof shape between the two cases, the contents would have had a greater chance of survival in the target case. He might adjust his vulnerability prediction to only 60% because of that factor.

The first four subjects were asked to compare their target case to three comparison cases. Data on the causal factors were listed with the photographs for each case. First, viewing photos of the comparison cases after blast, SMEs judged the vulnerability of specific contents that might have been inside them. Then they examined the target case and, by comparing it to the comparison cases, made a prediction of the vulnerability of the same contents in the target scenario. Finally, they were asked to specify the degree to which each causal factor contributed to the adjustment: for example, if a difference in roof shape caused them to decrease the predicted vulnerability, by how much did they lower the probability that the contents would have ceased to function.

This strategy proved to be both unwieldy and confusing. After analyzing the results of these interviews, it was clear that the SMEs considered factors in combination rather than singly. They could not treat the effects of changes in individual factors in an additive fashion, nor were they able to combine them in any satisfactory formula. They therefore could not with comfort quantify the adjustment made for each of the factors, although they were able to make over-all adjustments on the basis of the total scenario.

We therefore simplified the strategy, as described in this report. Subjects did not attempt to quantify the effect of differences in each causal factor. Rather, they made a global estimate of the probability of failure in the new case, compared to the comparison. Then they reviewed the causal factors they had considered in making the estimate, and rated each as to its importance and the direction of its influence on the prediction.

The Pilot Study did yield prediction scores (probabilities that equipment would cease to function), shown in Table C:1. These were not given any statistical analysis.

Table C:1
PREDICTION SCORES¹ OF PILOT STUDY SMEs

Target case	SME#	Initial Prediction	Post Prediction	Comparison Case	B4	B5	Prediction B6
A2	4	.70	---	.05	1.00	.50	
A3	1	.90	.85 - .90	.98	.40	.90 - .95	
A3	2	.99	.99	.99	.90	.99	
A3	3	.90	.85 - .90	.98	.40	.90 - .95	

¹ Prediction Scores are estimates of the probability that the specified contents of the structure will have ceased to function immediately following the blast.